

# (11) EP 0 933 423 B1

(12)

# **EUROPEAN PATENT SPECIFICATION**

- (45) Date of publication and mention of the grant of the patent: 22.08.2007 Bulletin 2007/34
- (21) Application number: 96935358.0
- (22) Date of filing: 17.10.1996

(51) Int Cl.:

C12N 15/57 (2006.01) C12N 15/12 (2006.01) C07K 14/435 (2006.01) C07K 16/18 (2006.01) C12P 21/02 (2006.01)

A61K 39/395 (2006.01)

C12N 9/64 (2006.01) C07K 14/47 (2006.01) C12P 21/08 (2006.01) C12N 1/21 (2006.01) A61K 38/17 (2006.01) A61K 31/70 (2006.01)

- (86) International application number: PCT/JP1996/003017
- (87) International publication number:WO 1997/031109 (28.08.1997 Gazette 1997/37)

(54) MELTRINS

**MELTRINE** 

**MELTRINES** 

- (84) Designated Contracting States:

  AT BE CH DE DK ES FI FR GB GR IE IT LI LU MC

  NL PT SE
- (30) Priority: 23.02.1996 JP 6175696
- (43) Date of publication of application: 04.08.1999 Bulletin 1999/31
- (83) Declaration under Rule 28(4) EPC (expert solution)
- (60) Divisional application: **07011948.2**
- (73) Proprietor: MOCHIDA PHARMACEUTICAL CO., LTD. Shinjuku-ku, Tokyo 160-0004 (JP)
- (72) Inventors:
  - FUJISAWA, Atsuko, Room 305
     Koganei-shi, Tokyo 184 (JP)
  - YAMAKAWA, Toru, Mochida Pharmaceutical Co., Ltd. Tokyo 160 (JP)
  - SHIRAKAWA, Kamon, Mochida Pharmaceutical Co., Ltd. Tokyo 160 (JP)

- MiZUSHIMA, Chitose, Mochida Pharmaceutical Co. Ltd Tokyo 160 (JP)
- OGAWA, Naoki, Mochida Pharmaceutical Co., Ltd. Tokyo 160 (JP)
- (74) Representative: Woods, Geoffrey Corlett
   J.A. KEMP & CO.
   Gray's Inn
   14 South Square
   London WC1R 5JJ (GB)
- (56) References cited:

WO-A-97/09420

WO-A-97/40072

- EMBL DATABASEAccession no D50411, Sequence Indentity MMMAB 6 November 1995 FUJISAWA SEHARA A: "Mouse mRNA for meltrin alpha" XP002151067
- EMBL DATABASEAccession no D50410, Sequence indentity MMMBA 6 November 1995 FUJISAWA SEHARA A: "Mouse mRNA for meltrin beta" XP002151068
- EMBL DATABASEAccession no D50412, Sequence indentity MMMGC 6 November 1995 FUJISAWA SEHARA A: "Mouse mRNA for meltrin gamma" XP002151069

EP 0 933 423 B1

Note: Within nine months from the publication of the mention of the grant of the European patent, any person may give notice to the European Patent Office of opposition to the European patent granted. Notice of opposition shall be filed in a written reasoned statement. It shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

- WESKAMP G ET AL: "MDC9, A WIDELY EXPRESSED CELLULAR DISINTEGRIN CONTAINING CYTOPLASMIC SH3 LIGAND DOMAINS" THE JOURNAL OF CELL BIOLOGY, US,ROCKEFELLER UNIVERSITY PRESS, vol. 132, no. 4, 1 February 1996 (1996-02-01), pages 717-726, XP000644308 ISSN: 0021-9525
- NOMURA N ET AL: "PREDICTION OF THE CODING SEQUENCES OF UNIDENTIFIED **HUMAN GENES. I. THE CODING SEQUENCES OF** 40 NEW GENES (KIAA0001-KIAA0040) **DEDUCEDBY ANALYSIS OF RANDOMLY** SAMPLED CDNA CLONES FROM HUMAN **IMMATURE MYELOID CELL LINE KG-1" DNA** RESEARCH, JP, UNIVERSAL ACADEMY PRESS, vol. 1, no. 1, 1994, pages 27-35, XP002049267 ISSN: 1340-2838 - & EMBL DATABASE Accession no D14665, Sequence Identity HSORF0931 March 1993 NOMURA N ET AL: Prediction of new human genes..... "XP002151070 -& SWISSPROT DATABASE Accession no Q13443 01 November 1996 "Cellular disintegrin-related protein precursor" XP002151071
- CELL STRUCTURE AND FUNCTION, 20(6), (1995),
   T. YAGAMI-HIROMASA et al., "Meltrinalpha, A Novel Metalloprotease-Disintegrin, Participates in Myotube Formation", p. 585 2A-1345.

- CELL STRUCTURE AND FUNCTION, 20(6), (1995),
   T. KURISAKI et al., "Cloning of Meltrins, Novel Mambers of Metalioprotease-DisIntegrin Family Involved in Myogenesis", p. 585 2A-1330.
- NATURE, 377, (1995), T. YAGAMI-HIROMASA et al., "A Metalloprotease-Disintegrin Participating in Myoblast Fusion", pages 652-656.
- GENERALIZED RESEARCH REPORT IN FISCAL YEARS HEISEI 5 TO 7, MORPHOLOGICAL AND BIOCHEMICAL AND MOLECULAR BIOLOGICAL BASIC RESEARCH ON MUSCULAR DYSTROPHY, (03.96), JUNKO FUJISAWA et al., "Molecular Mechanism of Differentiation and Formation of Skeletal Muscle", pages 85-88.
- GENOMICS, 34(3), (15.06.96), C. CHO et al.,
   "Chromosomal Assignment of Four Test is Expressed Mouse Genes from a New Family of Transmembrane Proteins (ADAMs) involved in Cell-Cell Adhesion and Fusion", pages 413-417.
- EXPERIMENTAL MEDICINE, 14(10), (01.07.96), TOMOHIRO KURIZAKI, JUNKO FUJISAWA (SEBARA), "A Novel Cellular Adhesive Molecule Meltrin Belonging to Metalloprotease Dysintegrain Family", pages 1352-1356.
- BIOCHEMISTRY,68(8), (08.96), TSUYOSHIENDO,
   "ADAM Family and Cell Fusion", pages 1453-1458
- PROC. NATL. ACAD. SCI. U.S.A., 91, (1994), G.
   WESKAMP et al., "A Family of Cellular Proteins Related to Snake Venom Disintegrins".

#### Description

### Technical Field

[0001] This invention relates to Meltrins and polypeptides of the respective domains thereof; DNAs encoding the same; antisense oligonucleotides for these DNAs; various antibodies against these Meltrins and the polypeptides of the respective domains thereof; expression vectors comprising the DNAs; transformants constructed by using these expression vectors; a process for producing the above-mentioned meltrins and the polypeptides of the respective domains thereof by means of the transformants; and medical compositions comprising the Meltrins or Meltrin antagonists as an effective ingredient.

## **Background Art**

[0002] In the course of myotube formation, myoblasts, which have divided from myogenic cells originating in undifferentiated mesodermal cells and grown to differentiate, will start synthesizing muscle-specific substances such as myosin and actin after its final division, and will lose cell boundaries at the fusion surface to be tansformed into multinucleate syncytium named myotube through adhesion and fusion of cytoplasmic membranes with neighbouring cells of the same kind.

[0003] There have been already reported several kinds of membrane proteins involved in the myotube formation, such as N-Cadherin (Knudsen, K.A. et al., Expl. Cell Res., 188, 175-184 (1990), Merge, R.M. et al., J. Cell Sci., 103, 897-906 (1992)), M-Cadherin (Donalies, M. et al., Proc. Natl. Acad. Sci., U.S.A. 88, 8024-8028 (1991)), N-CAMs (Merge, R.M. et al., J. Cell Sci., 103, 897-906 (1992) and others), V-CAMs and Integrins (Rosen, G.D. et al., Cell 69, 1107-1119 (1992) and others).

[0004] However, the molecular mechanism has not yet been sufficiently understood concerning the course of formation of the multinucleate syncytium named myotube through adhesion and fusion of the cytoplasmic membranes of the myoblasts with each other.

[0005] On the other hand, the substances named "fusion peptides" have been known as an adhesion factor involved in the course of infectior of cells with viruses (Morrison, T.G. Virus Res., 10, 113-136 (1988) and the others). Fertilin, which was recently isolated as a factor involved in sperm-egg adhesion, has been found to contain a sequence similar to the fusion peptide of rubella virus (Biobel, C.P. et al., Nature 356, 248-252 (1992) and the others).

[0006] Many substances having adhesion activity are known as mentioned above, and substances which may inhibit the activity of Integrins and the like have been developed and studied as potential medical agents.

[0007] The present inventors have now isolated novel substances involved in adhesion. Particularly, on the assumption that some fusion peptide-like adhesion factor like in sperm-egg adhesion may be involved in adhesion and fusion of the myoblasts with each other in the course of myotube formation, the novel substances involved in cell adhesion have been cloned and named "Meltrins", by using highly conserved sequences in Fertilin  $\alpha$  and  $\beta$  as a probe.

## Disclosure of Invention

[0008] The present invention relates to a novel "Meltrin." "Meltrins" are characterized as proteins which are expressed in the course of differentiation-induction of muscle cells and to contain the highly conserved sequences in Fertilin α and β. Meltrins are also characterized as proteins which are involved in fusion, adhesion, or aggregation of cells. Thus, some kinds of cells such as muscle ones may fuse, aggregate or adhere via Meltrins.

[0009] Cell fusion means that more than two cells fuse with each other to form one multicleate syncytium. Adhesion of cells means that more than two cells adhere to each other. Aggregation of cells means that more than two cells (particularly the cells present in liquid) flock together to form a mass of cells. It may be considered that cells adhere to each other, followed by cell fusion and aggregation.

[0010] According to the invention, there is thus provided a soluble meltrin polypeptide which does not comprise a transmembrane domain or an intracellular domain and which comprises the amino acid sequence of Gly (No.1) to Ile (No. 686) from the N-terminal in Fig. 15a - Fig. 15f, or the amino acid sequence of Glu (No. 156) to Ile (No. 686) from the N-terminal in Fig.15a - Fig.15f.

[0011] The invention also provides:

- a DNA comprising a base sequence encoding the polypeptide of the invention;
- a DNA of the invention which comprises the base sequence of No.1 to No. 2058 from the 5' terminal in Fig. 15a Fig. 15f;
  - a DNA of the invention which comprises the base sequence of No. 1 to No. 2848 from the 5' terminal in Fig. 15a Fig. 15f;

- an antisense oligonucleotide which hybridizes with a part of the sequence of No. 1957 to No. 2848 from the 5' terminal in Fig. 15a Fig. 15f;
- an antibody which recognizes the C-terminal region of a meltrin wherein the C-terminal region is from amino acid No. 653 to No. 686 from the amino terminal in Fig. 15a - Fig. 15f;
- a vector comprising a DNA of the invention;

5

- a transformant by the vector of the invention;
- a process for producing a polypeptide of the invention, which process comprises culturing the appropriate transformant of the invention:
- a medical composition comprising a polypeptide, an antisense oligonucleotide or an antibody of the invention;
- use of a polypeptide, an antisense oligonucleotide or an antibody of the invention for the manufacture of a medicament for treatment of a condition associated with unhealthy enhanced bone resorption; and
  - use of a polypeptide, an antisense oligonucleotide or an antibody of the invention for the manufacture of a medicament for preventing metastasis of cancer cells.
- At least three kinds of molecules  $(\alpha, \beta)$  and  $(\alpha, \beta)$  have been isolated from one animal species.
  - [0012] Meltrins may be mouse Meltrins  $\alpha$ ,  $\beta$  and  $\gamma$ , which are characterized by amino acid sequences shown in Fig. 2a Fig.2i, Fig.3a Fig.3j and Fig.4a Fig.4i, respectively, or partial sequences thereof.
  - [0013] Other examples are human Meltrins  $\alpha$ ,  $\beta$  and  $\gamma$ , which are characterized by amino acid sequences shown in any one of Fig.12a Fig.12b, Fig.15a Fig.15f or Fig.23a Fig.23b; any one of Fig.16 or Fig.17a Fig.17c; or Fig.13a Fig.13d, respectively, or partial sequences thereof.
  - [0014] The above amino acid sequences should be considered only examples of Meltrins. Any variant of the above amino acid sequences wherein a part of the sequences has changed due to deletion, substitution, addition, insertion and the like of amino acids is therefore a Meltrin, as long as it is expressed in muscle cells, and have the highly conserved sequences in Fertilin  $\alpha$  and  $\beta$  or is involved in fusion, adhesion or aggregation of cells. As shown now by the present inventors, a high homology is seen in the part from disintegrin domain to cysteine-rich region of mouse amino acid sequences shown in Fig.2a Fig.2j and human amino acid sequences shown in Fig.12a Fig.12b. It is considered that such substances as showing homology of about 80 % or more, preferably about 90 % or more to the above amino acid sequences may keep the function as Meltrin. Particularly, it is believed that the substances having the sequences with homology of about 80 % or more, preferably about 90 % or more to the region from metalloproteinase domain to disintegrin domain of mouse or human Meltrins  $\alpha$ ,  $\beta$  and  $\gamma$  will have substantially the same activity, even if all of the other sequences are different from them. Accordingly, Meltrins may include substances having a high homology to the above amino acid sequences or to a part thereof and showing substantially the same activity as mouse or human Meltrins.
  - [0015] In other words, Meltrins may be characterized by having amino acid sequences encoded by base sequences that may hybridize the sequences complementary to the base sequences encoding any one of the amino acids shown in Fig.2a Fig.2j, Fig.3a Fig.3j, Fig.4a Fig.4i, Fig.12a Fig.12b, Fig.13a Fig.13d, Fig.15a Fig.15f, Fig.16, Fig.17a Fig.17c or Fig.23a Fig.23b.
  - [0016] Meltrins exist in bodies as a membrane protein consisting of intracellular domain, transmembrane domain, and extracellular domain; and as a soluble protein having no transmembrane domain. The extracellular domain contains precursor domain, metalloproteinase domain, disintegrin domain, and cysteine-rich region. Meltrin  $\alpha$  has a fusion peptide-like sequence in its cysteine-rich region (Refer to Fig.8).
  - [0017] The disintegrin domain is indispensable for the function of Meltrins such as adhesion, fusion and aggregation of cells. On the other hand, the precursor and metalloproteinase domains are thought to be regulating sequences for Meltrins to show the activity in a specific organ or tissue, or under specific conditions. It is known that the disintegrin found in snake venom will adhere to platelet IIb/IIIa. It is therefore presumed that the disintegrin domain by itself may have the function to adhere to cells. The metalloproteinase domain may act by itself as a protease as such.
  - [0018] A polypeptide may comprise any part of a Meltrin. Such polypeptides include the respective domain per se of Meltrins, polypeptides comprising at least the respective domain of Meltrins, any part of the sequences of Meltrins, polypeptides comprising at least any part of the sequences of Meltrins, and polypeptides comprising at least the sequence having the combination of any of the respective domains of Meltrins and any part of Meltrins in any order.
- 50 [0019] The above polypeptides which are chemically modified or formed into salts thereof.
  - [0020] The preferable examples of the such polypeptides include polypeptides consisting of a part of the disintegrin domain, polypeptides consisting of the disintegrin domain per se, polypeptides comprising at least the disintegrin domain, polypeprides comprising at least the disintegrin and cysteine-rich regions, polypeprides comprising at least the metalloproteinase, disintegrin and cysteine-rich regions, polypeptides consisting of a part of the metalloproteinase domain, and polypeptides consisting of the metalloproteinase domain per se.
  - [0021] There may be mentioned as other preferable examples of the present polypeptides those comprising at least the disintegrin and cysteine-rich regions, but not comprising the transmembrane domain, or comprising neither the transmembrane domain nor intracellular domain; and those comprising at least the metalloproteinase, disintegrin and

cysteine-rich regions, but not comprising the transmembrane domain, or comprising neither the transmembrane domain nor intracellular domain. Such polypeptides comprising no transmembrane domain are a soluble one which will be secreted through a cell membrane into extracellular area. The soluble polypeptides may be collected from supernatant of the culture medium of cells. When optionally combined downstream of a suitable signal sequence and expressed by cells in a genetic engineering process, it will be secreted into the culture supernatant and advantageously collected therefrom with a high efficiency.

[0022] The amino acid sequences in Fig.2a - Fig.2j, Fig.3a - Fig.3j, Fig.4a - Fig.4i, Fig.12a - Fig. 12b, Fig.13a - Fig. 13d, Fig.15a - Fig.15f, Fig.16, Fig.17a - Fig.17c and Fig.23a - Fig.23b, which correspond to the precursor domain, metalloproteinase domain, disintegrin domain, cysteine-rich region, intracellular domain, and transmembrane domain of mouse and human Meltrins  $\alpha$ ,  $\beta$  and  $\gamma$ , are discussed in the Examples. It should be noted, however, that the polypeptides having the above corresponding amino acid sequences constitute only examples. That is to say, polypeptides essentially comprising the same amino acid sequences may also be prepared. Thus, the boundaries of each domain are not limited to those defined in the Examples. Polypeptides comprising the domains wherein the boundaries are shifted to N-, C-terminals or both by 1 to about 20 amino acids from the boundaries defined in the Examples may also be prepared, as long as they have substantially the same function as that of the above polypeptides. Similarly, the polypeptides wherein a part of the amino acid sequences has changed due to deletion, substitution, addition, insertion and the like of amino acids may also be prepared, as long as they have substantially the same function as that of each domain.

[0023] It is considered that the polypeptides comprising such amino acid sequences as showing homology of about 80 % or more, preferably about 90 % or more to the amino acid sequences in each domain of the above figures may have the same function as that of the polypeptide of the present invention.

[0024] The Meltrin of the present invention may be used to bond cells to each other or to apparatuses such as a plate. It may be also fused with any other substances to efficiently deliver the substances to muscle cells upon its application into culture systems of the muscle sells, tissues or bodies.

[0025] On the other hand, the polypeptides comprising at least a part of Meltrins may be added to the culture systems to competitively inhibit the adhesion, fusion or aggregation of cells. Particularly, the disintegrin domain per se, a part thereof, or a soluble polypeptide comprising the disintegrin domain may be used as an effective ingredient in a medical composition for inhibiting the adhesion of cells. For example, such medical composition may be used as an anticoagulant to inhibit thrombus formation or blood coagulation, and be used to treat thrombosis, DIC and multi-organ failure. Furthermore, since it is considered that adhesion factors such as integrin family are involved in metastasis of cancer cells, the polypeptides comprising the disintegrin domain may be used as a drug for inhibiting the growth of cancers, or the adhesion of cancer cells to other cells so as to prevent their metastasis. In addition to the above, it is known that the adhesion of cells plays an important role in formation of osteoclast. The examples will demonstrate that Meltrins are involved in the adhesion in the formation of osteoclast, and anti-Meltrin antibodies may inhibit the formation of osteoclast and the increase of bone resorption. Accordingly, the polypeptide of the present invention may be used as an effective ingredient in a medical composition for inhibiting the increase of bone resorption, like as anti-Meltrin antibodies,

[0026] Among the polypeptides comprising at least a part of the Meltrin of the present invention, those comprising the metalloproteinase domain may act as a protease by itself, or be used to competitively inhibit the activity of other proteases so that they may be utilized as a drug for treating inflammatory diseases.

[0027] The Meltrin polypeptide of the present invention may also be used as an antigen for producing antibodies.

[0028] The present invention also relates to DNAs comprising the base sequence encoding the amino acid sequences of the Meltrin of the present invention or the polypeptides comprising any parts thereof.

[0029] The above DNAs include any type of DNAs such as genomic DNAs and cDNAs.

20

35

[0030] Examples of DNAs encoding Meltrins are those encoding mouse Meltrins  $\alpha$ ,  $\beta$ , and  $\gamma$ , or the polypeptides comprising any parts thereof, which are characterized by the coding regions shown as the base sequences in Fig.5a - Fig.5j, Fig.6a - Fig.6h, and Fig.7a - Fig.7e, respectively, or partial sequences thereof. Other examples are those encoding human Meltims  $\alpha$ ,  $\beta$ , and  $\gamma$ , or the polypeptides comprising any parts thereof, which are characterized by the coding regions of the sequences shown as the base sequences in any one of Fig.12a - Fig.12b, Fig. 15a - Fig. 15f or Fig.23a - Fig.23b; any one of Fig.16 or Fig.17a - Fig.17c; or Fig.13a - Fig.13d, respectively, or partial sequences thereof.

[0031] The base sequences in the above figures, which correspond to the precursor domain, metalloproteinase domain, disintegrin domain, cystein-rich domain, intracellular domain, and transmembrane domain of mouse and human Meltrins  $\alpha$ ,  $\beta$  and  $\gamma$ , are discussed in the Examples. It should be noted, however, that they constitute only examples of such DNAs. DNAs essentially comprising the same base sequences may also be prepared.

[0032] Thus, the boundaries of each domain are not limited to those defined in the Examples. And the DNAs comprising sequences encoding the domains wherein the boundaries are shifted to 5'-, and/or 3'-ends by 1 to about 60 base pairs from the boundaries defined in the Examples may also be prepared, as long as they encode the polypeptides having substantially the same function as that of each domain.

[0033] In addition of the above base sequences, DNAs may be prepared comprising the base sequences or partial sequences thereof, which encode the same amino acid sequences as above prepared by means of chemical synthesis

or genetic engineering in consideration of degeneracy of codons. As now shown by the present inventors, a high homology is seen in mouse and human Meltrins. It is therefore considered that the substances showing homology of about 80 % or more, preferably about 90 % or more to the above amino acid sequences may keep the function as Meltrin, and that DNAs encoding such homologous polypeptides will hybridize with each other. Accordingly, DNA fragments may be obtained by hybridization under stringent conditions using the DNAs having the base sequences complementary to those in the above figures as a probe.

[0034] The DNAs of mouse or human Meltrins  $\alpha$ ,  $\beta$  and  $\gamma$ , or partial sequences thereof may be inserted into plasmid vectors. Strains of E. coli transformed by the same plasmid vectors have been deposited with the National Institute of Bioscience and Human-Technology, Agency of Industrial Science and Technology.

[0035] The DNAs described herein may be prepared by known methods. The cDNAs, for example, may be prepared by using cDNA library and known PCR (e.g., Michael A.I. et al., PCR Protocols, a guide to method and application, Academic Press, 1990) with degenerative primers for a part of the amino acid sequences (for example, the degenerative primer encoding the amino acid sequences of the disintegrin domain) shown in Fig.2a - Fig.2j, Fig.3a - Fig.3j, Fig.4a - Fig.4i, Fig.12a - Fig. 12b, Fig.13a - Fig.13d, Fig.15a - Fig.15f, Fig.16, Fig.17a - Fig.17c and Fig.23a - Fig.23b. The DNAs described herein may also be prepared by hybridization method using a probe prapared on the basis of the base sequences of the above amplified DNA fragments.

[0036] As demonstrated in the Examples, the preferable source of cDNA library include cells obtained by inducing myoblast to differentiate, bone marrow and fetal pulmonary cells. Known cDNA libraries prepared from placenta, chorionic cells and fetal cells may also serve as the source of cDNA library in the present invention.

[0037] Among the DNAs described herein, one encoding the polypeptide in which any parts of Meltrins are combined in any order may be prepared by the following steps. That is, each DNA fragment encoding any part of Meltrins is amplified by PCR, in which the primers may be optionally modified in order to provide an appropriate restriction enzyme site. The amplified DNA fragments are ligated with each other by DNA ligase, so that a reading frame should not be shifted.

[0038] The DNAs described herein may be used for producing the Meltrins or polypeptides of the present invention by means of genetic engineering. Such prodution may be carried out with reference to known methods (for example, Sambrook J. et al., Molecular Cloning a Laboratory Manual 2nd ed., Cold Spring Harbor Laboratory, New York, 1989).

[0039] The DNAs described herein inserted into suitable vectors may also be used in gene therapy. The base sequence encoding any physiologically active substances is fused downstream of the present DNAs followed by insertion of the resulting fused DNA into a vector originated in an appropriate virus, and cells in a living body are transformed with the resulting vector, so that the physiologically active substances may be expressed as a fused protein with the Meltrin of the present invention. The thus expressed physiologically active substances will be delievered near to the cells to which Meltring adhere.

[0040] The present invention further relates to antisense oligonucleotides and derivatives thereof for the DNAs encoding the Meltrin of the present invention or for the polypeptides comprising any part thereof.

[0041] The present antisense oligonucleotides and derivatives thereof are characterized by their base sequences complementary to those encoding Meltrins or a part thereof, or by their function to inhibit the expression of Meltrins or the polypeptides comprising any part thereof. The antisense oligonucleotides and derivatives thereof characterized by the latter feature include those complementarily bonding to the non-coding regions existing upstream or downstream of the coding regions of Meltrins as well as those complementarily bonding to the coding regions of Meltrins or any part thereof.

[0042] Examples of the antisense oligonucleotides and derivatives thereof described herein include the base sequences complementary to the DNAs of the present invention or any part thereof, particularly to those shown in Fig. 5a - Fig. 5j, Fig.6a - Fig.6h, Fig.7a - Fig.12a - Fig.12b, Fig.13a - Fig.13d, Fig. 15a - Fig. 15f, Fig.16, Fig.17a - Fig.17c and Fig.23a - Fig.23b. Uracil (U) may be used instead of thymine (T) as a complementary base to adenine (A).

[0043] The derivatives of the present antisense oligonucleotides include any one that is similar to the antisense oligonucleotides in steric structure and function, such as those wherein other substances are bound to 3'- or 5'-end of the oligonucleotides; those wherein at least one of bases, sugars or phosphoric acids in the oligonucleotides has substitution or modification; those having non-naturally occurring bases, sugars or phosphoric acids; and those having back bone other than that of sugars-phosphoric acids.

[0044] The antisense oligonucleotide of the invention and derivatives thereof may be prepared by known methods (for example, ed., Stanley T. Crooke and Bernald Lebleu, in Antisense Research and Applications, CRC Publishing, Florida, 1993).

[0045] The present antisense oligonucleotide of a naturally occurring type may be prepared by chemically synthesizing sense-primers and antisense-primsers having the base sequences complementary to 3'- or 5'-end of the antisense oligonucleotide sequences, followed by PCR using the Meltrin genes or RNAs encoding Meltrins as a template. Otherwise, the derivatives of the antisense oligonucleotides such as a methylphosphonate and phosphorothionate types may be prepared by means of a chemical synthesizer (e.g., Perkin Elmer Japan Co., Type 394) according to the manual attached to the chemical synthesizer, followed by, if necessary, purification of the synthesized products in HPLC method using

reversed phase chromatography and the like.

25

[0046] The present antisense oligonucleotide and derivatives thereof may be labelled with radioisotopes, fluorescent substances, enzymes or luminescent substances and used as a probe for detecting the existence of Meltrins or any part thereof in a sample. The present antisense oligonucleotide may also be used as a medical composition for inhibiting the expression of Meltrins in a living body.

[0047] For the purpose of inhibiting the expression of Meltrins by using the present antisense oligonucleotide and derivatives, they may be solubilized or suspended in a suitable solvent, enclosed in a liposome, or inserted into a suitable vector.

[0048] It is preferred that the present antisense oligonucleotide and derivatives thereof used in the medical composition should have a pharmaceutically acceptable purity and be used in a pharmaceutically acceptable way.

[0049] As already mentioned in the above, it is considered that Meltrins are involved in formation of osteoclast, growth and metastasis of cancers as well as skeletal myogenesis. Accordingly, the present antisense oligonucleotide and their derivatives which are capable of inhibiting the expression of Meltrins may be used in in the manufacture of a medicament treatment and prevention of cancers, treatment of osteoporosis and hypercalcemia by inhibiting bone resorption.

[0050] The present invention also relates to antibodies recognizing the Meltrin of the present invention or the polypeptides comprising at least any part thereof. In other words, they include those recognizing only Meltrins of the present invention, those recognizing only the polypeptides of the present invention and those recognizing both of them.

[0051] Antibodies described herein include those cross reacting with other polypeptides in addition to those specifically recognizing Meltrins and the polypeptides of the present invention. They also include those specifically recognizing any one of Meltrins  $\alpha$ ,  $\beta$  and  $\gamma$ , and those specifically recognizing more than two of Meltrins  $\alpha$ ,  $\beta$  and  $\gamma$ , as well as those recognizing only Meltrins originated in a particular animal such as human and mouse or only the polypeptides comprising at least any part thereof, and those recognizing Meltrins originated in more than two kinds of animals or the polypeptides comprising at least any part thereof.

[0052] Such antibodies include those recognizing the amino acid sequences in Fig.2a - Fig.2j, Fig.3a - Fig.3j, Fig.4a - Fig.4i, Fig.12a - Fig.12b, Fig.13a - Fig.13d, Fig.15a - Fig.15f, Fig.16, Fig.17a - Fig.17c or Fig.23a - Fig.23b, or any part thereof

[0053] More preferably, the antibodies described herein are those obtained by immunization of animals with the polypeptides comprising said amino acid sequences or any part thereof as an antigen, which may be optionally conjugated with a suitable carrier.

[0054] Such antibodies may be prepared by inserting DNA comprising the base sequnces shown in Fig.5a - Fig.5j, Fig.6a - Fig.6h, Fig.7a - Fig.7e, Fig.12a - Fig.12b, Fig.13a - Fig.13d, Fig.15a - Fig.15f, Fig.16, Fig.17a - Fig.17c or Fig. 23a - Fig.23b or any part thereof into a suitable expression vector, tranforming a suitable host cell by the vector to produce Meltrins, which are purified from cell bodies of the transformant or culture medium and administered as an antigen. The cell bodies per se of the transformant or any cells expressing Meltrins per se may be administered as an antigen. Such transformant or cells may express any one of Meltrins  $\alpha$ ,  $\beta$  and  $\gamma$ , or more than two kinds of them. The antibodies may be also prepared by chemically synthesizing the polypeptides having a part of the amino acid sequences of Meltrins, conjugating them with a carrier such as KLH (Keyhole Limpet Hemocyanin) and administering them as an antigen.

[0055] It is possible to prepare the present antibody that may recognize the whole of Meltrins even when the part of Meltrins is used as an antigen to be administered. It is also possible to prepare the present antibody that may recognize human Meltrins or any part thereof even when mouse Meltrins or any part thereof are used as an antigen to administered. [0056] The antibodies described herein include monoclonal and polyclonal ones, and may belong to any class or subclass.

[0057] The antibodies may be prepared according to known methods (e.g., "Meneki jikkenho (Laboratory manual of Immunology)" published by Japan Immunological Society). An example of the known methods will be described below. [0058] A suitable cell is transformed by an expression vector comprising the coding regions of the base sequences shown in Fig.5a - Fig.5j, Fig.6a - Fig.6h, Fig.7a - Fig.7e, Fig.12a - Fig.12b, Fig.13a - Fig.13d, Fig.15a - Fig.15f, Fig.16, Fig.17a - Fig.17c or Fig.23a - Fig.23b or any part thereof, and used as an antigen as such. Alternatively, Meltrins produced by the transformant are purified from cell bodies of the transformant or culture medium to be used as an antigen, or polypeptides consisting of amino acid sequences shown in the above figures are chemically synthesized, cojugated with a carrier such as KLH (Keyhole Limpet Hemocyanin) and purified to be used as an antigen.

[0059] Animals are inoculated with the antigen thus prepared, alone or together with a suitable adjuvant such as Freund's complete adjuvant (FCA) or Freund's incomplete adjuvant (FIA), subjected to boosting at two to four-week intervals. After boosting, the blood is drawn from the animals and antiserum is obtained therefrom. Animals to be immunized may be selected from rat, mouse, rabbit, sheep, horse, fowl, goat, pig, cattle and the like, depending on the kind of the antibody to be desired. Polyclonal antibodies may be obtained by purification of the antiserum by known methods such as salting-out, ion-exchange chromatography, affinity chromatography and any combination thereof.

[0060] Monoclonal antibodies may be prepared as follows. Antibody-producing cells such as spleen cells and lym-

phocytes are collected from the immunized animals, fused with myeloma and the like by known methods using polyethyleneglycol, Sendai virus, electrical pulse to give hybridomas. Clones which produce an antibody bonding to the Meltrin of the present invention are then selected and cultured. Monoclonal antibodies of the present invention are purified from the culture supernatant of the selected clones by known methods such as salting-out, ion-exchange chromatography, affinity chromatography and any combination thereof.

[0061] The present antibodies may be neutralizing antibodies, which inhibit the fusion, adhesion or aggregation of cells by Meltrins. The neutralizing antibodies of the present invention include those that can completely inhibit the activity of Meltrins, and those partially inhibit the same.

[0062] The neutralizing antibodies may be screened by adding antiserum or culture supernatant of the hybridomas to the culture system of Meltrin-expressing cells to evaluate the degree of inhibition of fusion or aggregation of cells. After the screening, the desired antibodies may be purified from the thus selected antiserum or culture supernatant of the hybridomas by the known methods.

[0063] The antibodies of the present invention include Fab, F(ab'), F(ab'), and Fv, as long as they recognize and bond to the present polypeptides or any part thereof. A single chain Fv may be also included in the present antibodies, which is obtained by constructing a gene encoding the single chain Fv wherein H and L chains are linked into a single chain and being expressed by a suitable host cell. Chimera antibodies, human antibodies and humanized antibodies are also included in the present invention, as long as they recognize and bond to the present polypeptides or any part thereof. [0064] For example, the chimera antibodies may be prepared by substituting a gene encoding the constant region of human antibodies for a gene encoding the constant region of the mouse antibodies recognizing Meltrins or the polypeptides of the present invention, expressing the thus reconstituted gene in animal cells. The human antibodies may be prepared by, for example, in vitro sensitization method (Borrebaeck, C.A.K.J. Immunol., Meth., 123, 157, 1989) or the method using SCID mouse (Toshio KUDO, Tissue Culture, 19, 61-65, 1993). The humanized antibodies may be prepared by reconstituting a gene so that complementary determining regions (CDR) of the human antibodies are replaced with those of the mouse antibodies, and expressing the gene in animal cells (Carter et al., Pro. Nat. Acad. Sci, 89, 4285, 1992). [0065] If necessary, amino acids in a framework of the variable region of the humanized antibodies thus reconstituted may be replaced, so that the framework should have a high homology to that of the mouse antibodies and CDR of said humanized antibodies may form an appropriate antigen-binding site. The preferred examples of the humanized antibodies are those having the same CDR as the neutralizing antibodies F932-15-2 and F937-9-2. For the preparation of these preferred humanized antibodies, the DNA encoding the antibodies is prepared from the hyridoma F932-15-2 or F937-9-2, and linked with the DNAs encoding human antibodies so that the sequences other than CDRs should originate in the human antibodies. Any variation may be optionally introduced into the DNA encoding the framework portion. The thus obtained DNA is then inserted into a suitable expression vector to transform a suitable cell, and the humanized antibodies are purified from the culture supernatant of the transformant.

[0066] The present antibodies may be labelled with fluorescent substances, enzymes, luminescent substances or radioisotopes to detect Meltrins or their decomposed products present in body fluid or tissues. Since it is considered that Meltrins are involved in formation of myotube, resorption of bone and metastasis of cancers as already mentioned in the above, the detection of the existence of Meltrins in body fluid or tissues would make it possible to estimate the progress of diseases and prognosis and to confirm the effects of treatments. The present antibodies may be also used to provide an antibody affinity column, or to detect Meltrins in a fraction during the course of purification of Meltrins.

[0067] The neutralizing antibodies of the present invention may serve as an effective ingredient of a medical composition for inhibiting bone resorption, inflammatory diseases, blood coagulation and metastasis of cancers, owing to their ability to inhibit fusion or adhesion of cells. They may serve as an agent used in culture to inhibit the aggregation of cultured cells. When used as the effective ingredient of the medical composition, the human or humanized antibodies are preferred from the viewpoint of their antigenicity.

[0068] Also, the present invention relates to a vector comprising the DNA of the present invention. The present vector may further contain, if necessary, an enhancer sequence, promoter sequence, ribosome-binding sequence, base sequence for amplification of the number of copies, sequence encoding signal peptides, sequences encoding other polypeptides, poly(A)-additional sequence, splicing sequence, origin of replication, base sequence of the gene for selective markers and so on.

[0069] The present vector may be prepared by inserting the DNAs of the present invention into any vectors according to known methods (e.g., Molecular Cloning, a Laboratory Manual 2nd ed., Cold Spring Harbor Laboratory, New York, 1989). The preferable examples of the DNAs encoding Meltrins or any part thereof have been already disclosed in the present specification. The present vectors include a plasmid vector, phage vector and virus vector; pUC118, pBR322, pSV2-dhfr, pBluescriptII, PHIL-S1, λZap II, λgt10, pAc700, YRP17, pEF-BOS and pEFN-II being preferred.

[0070] The preferred vectors of the present invention may optionally comprise the origin of replication, selective markers, and promoter in addition to the DNAs encoding Meltrins or the polypeptides comprising at least any part thereof so as to be used to express Meltrins or the same polypeptides. As the origin of replication, ColEl, R factor, F factor and so on may be used in the vectors for E.coli; SV40- or adenovirus-derived ones in the vectors for animal cells; and ARS1-

derived one in the vectors for yeast. As the promoter, trp, lac and tac promoters may be used in the vectors for E. coli; SV40-, cytomegalovirus-, and adenovirus-derived ones, and those intrinsically existing in the genes of human or animals such as the promoter region of an elongation factor  $1\alpha$  in the vectors for animal cells; and  $\alpha$  promoter in the vectors for yeast, especially AOX1 promoter in the case of Pichia yeast. In the addition to the above sequences, the present vectors may further comprise, if necessary, RNA splicing site, signal for poly-adenylation and the like for the transforamtion of eucaryotic cells. The present vectors may be used for the production of Meltrins or any part thereof by means of genetic engineering, and used in gene therapy for Meltrins-related diseases.

[0071] The present invention therefore relates to transformants transformed by the above vectors.

[0072] The present transformants may be prepared by transforming suitable host cells by the above vectors according to known methods (e.g., Idenshi Kogaku Handbook (Handbook of gene technology), extra edition of Jikkenigaku, Yodo, 1991)). The host cells may be selected from procaryotic ones such as E.coli and Bacillus, or eucaryotic cells such as yeast, insect cells, and animal ones. The preferred transformants of the present invention are those derived from E.coli, yeast or CHO cell as a host cell to express Meltrins or the polypeptides of the present invention.

[0073] The present invention further relates to a process for producing Meltrins or the present polypeptides comprising at least any part thereof, comprising the step of culturing the above transformants.

[0074] In the present producing process, the transformants of the present invention are cultured, optionally with amplification of the gene or expression-induction, if necessary, according to known methods (e.g., Biseibutsugaku Jikkenho (Laboratory manual of microbiology), Tokyo Kagaku Dojin, 1992). The culture mixture, i.e., the cells and culture supernatant, is collected and optionally subjected to concentration, solubilization, dialysis, and various chromatography to purify Meltrins or the present polypeptides comprising any part thereof. The purification of the present polypeptides may be carried out by an optional combination of the above known methods for the purification of proteins, and an efficient purification could be performed by using an affinity column with the antibodies of the present invention.

[0075] In the present producing process, the polypeptides of the present invention may be produced by the transformants as a fused protein with other proteins such as  $\beta$ -galactosidase. In such case, the fused protein should be treated with chemicals such as cyanogen bromide or enzymes such as protease in a certain step in the purification process, so that the polypeptides of the present invention may be excised.

25

35

[0076] The present invention relates to medical compositions comprising a novel effective ingredient, which is Meltrins of the present invention or Meltrin-antagonist. The "Meltrin-antagonist" means a molecule which is able to inhibit fusion, adhesion or aggregation of cells through Meltrins. It includes, for example, the present antibodies recognizing Meltrins and having a neutralizing activity, the fragments of the same antibodies, the polypeptides consisting of any part of Meltrins or any combination thereof in any order, the antisense oligonucleotides for the DNAs encoding Meltrins or derivatives thereof.

[0077] The antibodies recognizing Meltrins may be prepared by the methods already mentioned in the above, and from which the antibodies which may completely or partially neutralize fusion, adhesion or aggregation of muscle cells, osteoclast or cancer cells are selected and used as the effective ingredient of the present medical composistions. The antibodies to be used as the effective ingredient include those prepared by administering any polypeptides as the antigen into any animals, as long as they may recognize human Meltrins and inhibit fusion, adhesion or aggregation of human muscle cells, osteoclast or cancer cells. They may be polyclonal or monoclonal ones, being preferably the human or humanized antibodies, considering the fact that the medical compositions will be administered to human. The human or humanized antibodies may be prepared according to the methods already described in the above.

[0078] The above fragments to be used as the effective ingredient in the present medical compositions include Fab, F(ab'), F(ab'), and F(ab'), and F(ab') and F(ab')

[0079] The polypeptides having any part of Meltrins or any combination thereof in any order may be used as the effective ingredient of the medical compositions, as long as they have the activity of inhibiting fusion, adhesion or aggregation of cells.

[0080] The preferable examples of the above polypeptides include those comprising a part or the whole of the disintegrin domain of Meltrins, those comprising the metalloproteinase, disintegrin and cysteine-rich regions of Meltrins, those comprising the disintegrin domain, but not comprising the transmembrane domain of Meltrins, and those comprising at least the metalloproteinase and disintegrin domains, but not comprising the transmembrane domain of Meltrins. These polypeptides may be chemically synthesized or produced by means of genetic engineering, as already mentioned in the above.

[0081] The antisense oligonucleotides or derivatives thereof to be used as the effective ingredient of the medical compositions may have any base sequences or any structure, as long as they are suitable for administration to human, and will complementarily bond to the gene for Meltrins to completely or partially inhibit their expression.

[0082] As already mentioned, Meltrins are involved in formation of osteoclast and metastasis of cancer cells. Accordingly, the medical comosition comprising the Meltrin-antagonist as the effective ingredient may be used for the purpose of inhibition of bone resorption or metastasis of cancers. The antagonist against human Meltrin  $\alpha$  or  $\beta$  is more preferably used as the effective ingredient in the medical composition for inhibition of bone resorption, while the antagonist against

human Meltrin  $\gamma$  is more preferably used as the effective ingredient in the medical composition for inhibition of cancer metastasis.

[0083] The Meltrins or Meltrin antagonist used as the effective ingredient in the present medical composition may be formed into their salts or be modified with pharmaceutically acceptable chemical agents, as long as they will never lose their essential activities. There may be exemplified as the salts those with inorganic acids such as hydrochloric acid, phosphoric acid, hydrobromic acid and sulfuric acid; those with organic acids such as maleic acid, succinic acid, malic acid and tartaric acid.

[0084] The medical compositions of the present invention include those administered by any route such as oral, subcutaneous, intravenous, intravenou

[0085] Any administration methods and intervals may be adopted. The present medical comopsitions may comprise depending on the administration route pharmaceutically acceptable auxiliaries such as fillers, packing agents, thickeners, binding agents, humidifying agents, disintegrating agents, surfactants, solution aids, buffers, pain-easing agents, preservatives and stabilizers. In the case of injections, for example, they may comprise stabilizers such as gelatin, human serum albumin (HSA) and polyethylene glycol; alcohols and saccharides such as D-mannitol, D-sorbitol, and glucose; and surfactants such as Polysorbate 80 (TM).

[0086] The medical compositions of the present invention may be mainly used for the prevention and treatment of osteoporosis and hypercalcemia, or the prevention of infiltration and metastasis of cancers.

[0087] The present medical compositions may be administered in an amount of about 0.1 - 100 mg/kg/day, preferably of about 1 - 50 mg/kg/day, more preferably of about 1 - 10 mg/kg/day, depending on the conditions or ages of patients, or administration routes. It may also be continuously administered by an intravenous drip, or administered by a single dose or doses at appropriate intervals per day.

[0088] The present medical compositions may be formulated according to the conventional manners. The injection, for example, may be formulated by dissolving the Meltrins or their antagonists aseptically prepared to a pharmaceutically acceptable purity into physiological saline, buffers and the like, followed by addition of gelatin or HSA, if necessary. Such injections may also be lyophilized, which will be dissolved into distilled water for the injections, physiological saline and the like when they are used.

[0089] The screening of the substances which may bind to Meltrins, inhibit the activity of Meltrins or regulate their expression may be carried out by using the Meltrins, various polypeptides, DNAs encoding them and the like.

## 30 Brief Description of Drawing

## [0090]

20

35

40

Fig.1a - Fig.1b show the comparison between parts of mouse Meltrins  $\alpha$ ,  $\beta$ ,  $\gamma$  (referred to as "M $\alpha$ ", "M $\beta$ ", "M $\gamma$ ") and the known sequences (macrophage specific antigen (MS2), Jararhagin (JR), fertilin- $\alpha$  (f $\alpha$ ).

Fig.2a - Fig.2j show the amino acid sequence of mouse Meltrin  $\alpha$  and its corresponding DNA sequence.

Fig.3a - Fig.3j show the amino acid sequence of mouse Meltrin  $\beta$  and its corresponding DNA sequence, wherein "N" means unidentified base.

Fig.4a - Fig.4i show the amino acid sequence of mouse Meltrin γ and its corresponding DNA sequence. "N" means unidentified base.

Fig.5a - Fig.5j show the result of DNA sequence analysis of the DNA inserted into pBSMel $\alpha$ , which comprises the base sequence encoding mouse Meltrin  $\alpha$ . "N", "M", "W" and "S" mean unidentified bases.

Fig.6a - Fig.6h show the result of DNA sequence analysis of the DNA inserted into pBSMelβ, which comprises the base sequence encoding mouse Meltrin β. "N", "M", "W" and "S" mean unidentified bases.

Fig.7a - Fig.7e show the result of DNA sequence analysis of the DNA inserted into pBSMelγ, which comprises the base sequence encoding mouse Meltrin γ. "N", "M", "W" and "S" mean unidentified bases.

Fig.8 shows schematically the structures of Meltrins  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\delta$ MP,  $\delta$ Pro.

Fig. 9 is a photograph of electrophoresis showing the result of Western blotting.

Fig. 10 is a photograph of electrophoresis showing the result of Northern blotting.

Fig.11a- Fig.11b show fusion-promoting activity of Meltrins for myoblast.

Fig.12a - Fig.12b show the result of base sequence analysis of the DNA inserted into pBShuMa300, which encodes human Meltrin  $\alpha$ . "N" and "X" mean unidentified bases and unidentified amino acids, respectively.

Fig. 13a - Fig. 13d show the result of base sequence analysis of the DNA inserted into pBShuM $\gamma$ G238, which encodes human Meltrin  $\gamma$ .

Fig. 14a shows schematically the cloning region in the cloning of human Meltrin  $\alpha$ .

Fig. 14b shows schematically the cloning region in the cloning of human Meltrin  $\beta$ .

Fig.15a - Fig.15f show partial amino acid sequence and its corresponding base sequence of human Meltrin  $\alpha$ , determined based on the result of analysis of the DNA inserted into pMel $\alpha$ -26N, pMel $\alpha$ -25C.

Fig.16 shows amino acid sequence and its corresponding base sequence of human Meltrin  $\beta$ .

Fig.17a - Fig.17c show partial amino acid sequence and its corresponding base sequence of human Meltrin  $\beta$ , determined based on the result of analysis of the DNA inserted into pMel $\beta$ -24C, pMel $\beta$ -24N.

Fig.18a shows schematically the sites of the peptides administered as the antigens in mouse Meltrin  $\alpha$ .

Fig.18b shows amino acid sequences of the peptides administered as the antigens.

Fig. 19 is a photograph of electrophoresis showing the result of Western blotting with anti-mouse Meltrin  $\alpha$  antibodies. Fig. 20 is a graph showing the inhibition of myotube formation by anti-mouse Meltrin antibodies.

Pig.21 is a graph showing the effects by anti-mouse Meltrin antibodies on the formation of pit (bone-resorption area) by mouse all bone cells.

Fig.22 is a graph showing the effects on the serum Ca values of the mouse fed with low Ca-content feed by antimouse Meltrin antibodies.

Fig.23a - Fig.23b show the amino acid sequence comprising the transmembrane domain of human Meltrin  $\alpha$  and its corresponding base sequence.

Fig.24a - Fig.24e show the result of base sequence analysis of the DNA inserted into pMelβ-24C, pMelβ-24N.

## Best Mode for Carrying Out the Invention

[0091] The present invention will be further illustrated by the following Examples, which should not be construed to limit the scope of the present invention.

### Examples

5

10

15

20

[0092] The abbreviations used in the following description are based on the conventional ones in the art.

[0093] The processes used in the following Examples are based on Sambrook J. et al., Molecular Cloning, a Laboratory Manual 2nd ed., Cold Spring Harbor Laboratory, New York, 1989; E. Harlow, D.Lane et al., Antibodies, A Laboratory Manual, Cold Spring Harbor Laboratory; and the like.

#### Example 1. Acquisition of the DNAs encoding mouse Meltrins by RT-PCR

## 30 (1) Preparation of RNA, cDNA.

[0094] Myogenic cell line derived from fetal fibroblast C3H10T1/2, (a clone transfected by the gene encoding "myogenin", a muscle differentiation-controlling factor and expressing the myogenin) was proliferated to the extent of 10<sup>6</sup> cells/φ 10 cm plate in DMEM supplemented with 10% fetal bovine serum (MOREGATE) and cultured at 37°C for 2 days in differentiation medium (DMEM containing 2 % horse serum from GIBCO) for differentiation and induction. Total RNA was separated according to Guanidine isothiocyanate/acid phenol method (Chomczynski P. and Sacchi N., Anal. Biochem., 162, 156-159, 1987), and poly (A) RNA was selectively separated by repeating twice oligo(dT)-cellulose column chromatography. By using the poly(A) RNA as a template and random primers (N6, Pharmacia), cDNAs were synthesized with MLV reverse transcriptase (GIBCO BRL) according to its manual for synthesis. The obtained cDNAs were then used as a template for the next PCR, and double strand DNAs were synthesized and inserted into a phage (λZapli (stratagene)) to give a cDNA library.

### (2) RT-PCR

[0095] RT-PCR was carried out by using the cDNAs prepared in the above (1) as a template in the following steps:
[0096] A degenerative primer encoding the amino acid sequence EDCDCG or EECDCG was synthesized and used as a sense primer, and a degenerative primer encoding the amino acid sequence KCGKLIC was synthesized and used as an antisense primer.

[0097] The primers were mixed with the above cDNAs, Taq polymerase and the reaction agents (Boehringer Manheim), and subjected to 36 reaction cycles of 95°C for 1 min, 55°C for 2 min, and 72°C for 3 min. The amplification product of around 450 bp was then collected by 1.5% agarose gel electrophoresis.

[0098] The amplified fragments thus obtained were inserted into a Smal site in the plasmid pBS-SKII(-) (stratagene), and subjected to DNA sequence analysis by means of a DNA sequencer (370A type, Applied Biosystems). As a result, it was found that three kinds of molecules (DNA fragments) existed (Fig.1), which were then used as a probe to screen the cDNA library so as to isolate cDNAs comprising an open reading frame with 903, 920 and 845 amino acid residues, respectively (Fig.2a - Fig.2j, Fig.3a - Fig.3j, Fig.4a - Fig.4i). The products of the respective genes were named Meltrins  $\alpha$ ,  $\beta$ , and  $\gamma$  (Fig.5a - Fig.5j, Fig.6a - Fig.6h, Fig.7a - Fig.7e). These cDNAs were inserted into pBS-SKII(-) to give the plasmids, "pBSMel $\alpha$ ", "pBSMel $\beta$ ", and "pBSMel $\gamma$ ", respectively.

[0099] E.coli strain JM109 was transformed according to a known method by the above plasmids "pBSMelα", "pBS-Melβ", and "pBSMelγ", respectively, and the resulting transformants "JM109(pBSMelα)", "JM109(pBSMelβ)", and "JM109 (pBSMelβ)" were deposited in the National Institute of Bioscience and Human-Technology, Agency of Industrial Science and Technology (1-3, Higashi 1-chome, Tsukuba-shi, Ibaraki-ken 350 Japan) on February 19, 1996 under accession numbers FERM P-15451, FERM P-15452, and FERM P-15453, respectively, and then transferred on October 8, 1996 to the deposit under the terms of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure and Regulation under accession numbers FERM BP-5701, FERM BP-5702, and FERM BP-5703, respectively.

#### (3) Analysis of the structure of Meltrins

10

20

40

[0100] From the structure analysis of Meltrins on the basis of the DNA sequences determined in the above (2), it was supposed that Meltrins  $\alpha$ ,  $\beta$ , and  $\gamma$  were a transmembrane-type protein consisted of an extracellular domain, transmembrane (TM) domain, and intracellular domain, and that the extracellular domain consists of a precursor domain (pro region) comprising a signal peptide-like sequence, metalloproteinase domain, disintegrin domain, and the following cysteine-rich region. A fusion peptide-like sequence was contained in the cysteine-rich domain of Meltrin  $\alpha$  (Fig.8).

[0101] Based on their homology to the snake venom, Jararhagin, it has been considered that in Meltrin  $\alpha$ , the precursor domain corresponded to the sequence from N-terminal to Arg (No.205) and to the bases No.221-835, the metalloproteinase domain to the sequence from Glu (No.206) to Pro (No.414) and to the bases No.836-1462, the disintegrin domain to the sequence from Phe (No.420) to Gly (No.509) and to the bases No.1478-1747, the cysteine-rich region to the sequence from His (No.510) to Gly (No.706) and to the bases No.1748-2338, the fusion peptide-like sequence to the sequence from Gly (No.585) to Glu (No.607) and to the bases No.1973-2041, the transmembrane domain to the sequence from Leu (No.707) to Leu (No.727) and to the bases No.2339-2401.

[0102] Similarly, it was considered that in Meltrin β, the precursor domain corresponded to the sequence from N-terminal to Arg (No.204) and to the bases No.63-674, the metalloproteinase domain to the sequence from Glu (No.205) to Pro (No.409) and to the bases No.675-1289, the disintegrin domain to the sequence from Tyr (No.415) to Gly (No. 504) and to the bases No.1305-1574, the cysteine-rich region to the sequence from Thr (No.505) to Pro (No.706) and to the bases No.1575-2180, the transmembrane domain to the sequence from Val (No.707) to Arg (No.729) or to Leu (No.724) and to the bases No.2181-2249 or 2181-2234.

[0103] Similarly, it was considered that in Meltrin  $\gamma$ , the precursor domain corresponded to the sequence from N-terminal to Arg (No.205) and to the bases No.69-683, the metalloproteinase domain to the sequence from Ala (No.206) to Pro (No.406) and to the bases No.684-1292, the disintegrin domain to the sequence from Tyr (No.412) to Gly (No. 502) and to the bases No.1302-1574, the cysteine-rich region to the sequence from Tyr (No.503) to Ala (No.694) and to the bases No.1575-2150, the transmembrane domain to the sequence from Leu (No.695) to Ile (No.714) and to the bases No.2151-2210.

### Example 2: Establishment of anti-Meltrin $\alpha$ antibodies

## (1) Preparation of immunogen

[0104] A chimera polypeptide was prepared as follows, which consisted of glutathione-S-transferase (GST) (Smith, D.B. & Johnson, K.S., Gene, Vol.67, 31-40, 1988) and the polypeptide having the amino acid sequence from Ser (No. 483) to Lys (No.635) of Meltrin  $\alpha$  in Fig.2a - Fig.2j, said polypeptide being attached to the C-terminal of GST. First, the plasmid, pGEX2T (Pharmacia) comprising the cDNA encoding GST was digested at a BamHI site and used as a vector. On the other hand, the cDNA corresponding to the amino acid sequence from Ser (No.483) to Lys (No.635) of Meltrin  $\alpha$  in Fig.2a - Fig.2j was amplified from pBSMel $\alpha$  by PCR, and ligated with a BamHI linker by a DNA ligase. The resulting cDNA was then ligated with the above vector by a DNA ligase to give a plasmid, which was then tranformed into E.coli strain NM522.

[0105] The transformed E.coli was cultured in L-broth with 1mM IPTG to produce a large amount of the chimera polypeptide in the inclusion bodies upon expression-induction. The strain was suspended into MTPBS (150mM NaCl, 16mM Na<sub>2</sub>HPO<sub>4</sub>, 4mM NaH<sub>2</sub>PO<sub>4</sub>, 0.1mM PMSF), subjected to ultrasonication, and solubilized with 1% Triton. The supernatant of the thus treated mixture was collected. Glutathione agarose (Sigma) was mixed with the supernatant to adsorb the chimera polypeptide which was then eluted with an elution buffer (50mM Tris-HCl, pH 8.0, 0.5mM glutathione) and used as an immunogen.

## (2) Preparation of antiserum

[0106] The antigen (1mg) prepared in the above (1) in 0.5ml PBS and RIBI in PBS 0.5ml (MPL+TDM+CWS Emulsion,

Funakoshi) was mixed with each other, and subcutaneously or intracutaneously administered into a rabbit (12 weeks old, female). After boosting three times with 500µg dose at 4 week intervals, the blood was collected and serum was separated to give antiserum.

### 5 (3) Affinity purification of antiserum

[0107] The chimera polypeptide expressed in E.coli and solubilized in the above (1), or GST having no fused polypeptide was bound to the glutathione agarose beads. The resulting beads were washed with 0.2M sodium borate (pH 9.0), and mixed with dimethyl pimelidiate (a final concentration of 20mM) so that the antigen was irreversibly bound to the beads, so as to give chimera polypeptide-affinity beads and GST-affinity beads, respectively.

[0108] The antiserum diluted ten times with 10mM Tris-HCl (pH 7.5) was first mixed with the GST-affinity beads for anti-GST antibodies to be absorbed and removed, and then mixed with chimera polypeptide-affinity beads for anti-Meltrin  $\alpha$  antibodies to be adsorbed thereon. The resulting chimera polypeptide-affinity beads were washed with 10mM Tris (pH 7.5) and 500mM NaCl, and the anti-Meltrin  $\alpha$  antibodies were eluted with 100mM glycine and collected as purified anti-Meltrin  $\alpha$  antibodies.

## (4) Western blotting

[0109] C2 cell was proliferated to the extent of 10<sup>6</sup> cells/\$\phi\$ 10cm plate in DMEM supplemented with 15% fetal bovine serum, then cultured at 37°C in differentiation medium (DMEM supplemented with 2% horse serum) and collected on the second day (C2DM d2) and on the 4th day (C2DM d4).

[0110] Further, C2 cell transformed by pBOSMelα (+) prepared in the following Example 5 (3) was cultured in DMEM supplemented with 15% fetal bovine serum at 31°C for three days, inoculated into a plastic dish (φ 6cm) at a density of 2 x 10<sup>5</sup>/dish, further cultured for one day and transferred into the above differentiation medium for differentiation induction. After two day-culture in the differentiation medium, the cells were collected.

[0111] The collected C2DM d2, C2DM d4 or transformants by pBOSMela (+) was mixed with SDS solubilizing buffer (100mM Tris-HCI (pH 6.8), 4% SDS, 20% Glycerol), subjected to ultrasonication and centrifuged to give their supernatant as a sample.

[0112] A membrane was wahsed twice with a washing solution. The antiserum prepared in the above (3) was diluted 20 times with 5% skim milk solution in TBS-T, into which the membrane was soaked and incubated at 37°C for one hour. After the incubation, the membrane was washed twice with the washing solution. The membrane was then soaked into a biotin-labelled anti-rabbit immunoglobulin antibody (Daco) diluted 4,000 times with the above skim milk solution and incubated at 37°C for one hour. After the incubation, the membrane was washed twice with the washing solution. The membrane was reacted with a peroxidase-labelled streptoavidin for one hour, washed twice, stained with MB reagent (Cat.TM912, Shic) and detected by ECL system (Amersham).

[0113] The results are shown in Fig.9.

[0114] The Western blotting revealed the bands at about 115KD, 86KD, 67KD, and 58KD, indicating that Meltrin  $\alpha$  was expressed as a glycoprotein. It was also considered that the precursor domain was deleted in the molecule of 86KD, and both the precursor and metalloproteinase domains were deleted in the molecule of 67KD or 56KD.

## Example 3: Northern blotting

40

[0115] Poly (A)+ RNAs were prepared from vaious tissues of mouse (bone, brain, liver, heart and skeletal muscle of adult mouse; bone and skeletal muscle of newborn mouse; and bone and skeletal muscle of fetal mouse) by using a mRNA purification kit of Pharmacia according to the method described in Example 1. RNAs were denatured by heating at 65°C for 5 min in 50% formamide, subjected to electrophoresis on 1.5% agarose gel comprising 6.6% formalin, and transferred onto a nylon membrane (Highbond-N, Amersham).

[0116] On the other hand, cDNAs encoding a part of the disintegrin and cysteine-rich regions (Glu(No.434) - Cys(No. 583) in Fig.2a - Fig.2j, Glu(No.429) - Cys(No.578) in Fig.3a - Fig.3j, Glu(No.426) - Cys(No.575) in Fig.4a - Fig.4i) were prepared by PCR, and labelled with <sup>32</sup>P using a random primer labelling kit (Megaprime, Amersham). As a control probe, cDNA encoding G3PDH (glyceraldehyde 3-phosphate dehydrogenase) was also llabelled with <sup>32</sup>P in the same way. The above mRNAs were hybridized with the radiolabelled cDNAs under high stringency conditions according to the method of Sambrook J.et al. (Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Habor Laboratory, New York, 1989).

<sup>55</sup> [0117] Their results are shown in Fig. 10.

[0118] Fig.10 has revealed that Meltrin  $\alpha$  and  $\beta$  were expressed only in bones of adult and newborn mouses, and skeletal muscles of newborn and fetal mouses (the results from the fetal mouse are not shown in Fgi.10). There was no tissue-specificity in the expression of Meltrin  $\gamma$ , since it was universally expressed in all the tissues.

#### Example 4: Confirmation of adhering activity of Meltrin $\alpha$

(1) Construction of plasmids pBOSMelαδMP(+) and pBOSMelαδMP(-)

6 [0119] A deletion type Meltrin δMP wherein the precursor and metalloproteinase domains in the extracellualr domain of Meltrin α had been deleted was prepared in the followig method.

[0120] The plasmid, pBSMel $\alpha$  was partially digested at MscI and subjected to electrophresis on 1% agarose gel to give a linear plasmid DNA. The resulting DNA was partially digested at Nhel, treated with a Klenow fragment to generate blunt ends, and subjected to intramolecular ligation. Vectors having the right deletion were selected and their DNA sequences were confirmed. After digestion at multicloning sites of EcoRV and Notl in the vectors, a deletion type  $\delta$ MP fragment of about 5.8kb was obtained.

[0121] On the other hand, the plasmid, pEFBOS (Mizushima S. & Nagata S, Nucleic Acid Res. Vol.18, p.5322, 1990) was digested by a restriction enzyme Xbal, dephosphorylated, treated with a Klenow fragment to generate blunt ends and subjected to electrophresis on 1% agarose gel to give a linear plasmid DNA. The resulting linear DNA was then ligated with the above fragment of about 5.8kb by a DNA ligase to give the plasmids pBOSMelαδMP(+) and pBOSMelαδMP(-). They were the constructs comprising the inserted DNA encoding the δMP fragment wherein the amino acid sequence of from lle(55) to Glu(399) of Meltrin α was deleted, in sense direction and antisense direction, respectively.

(2) Construction of plasmid pBOSMelα(+)

20

25

55

[0122] The plasmid, pBSMelα, was partially digested by EcoRV and Notl to give a fragment of about 7kb. The above pEFBOS plasmid was digested by a restriction enzyme Xbal, dephosphorylated, treated with a Klenow fragment to generate blunt ends, and subjected to electrophresis on 1% agarose gel to give a linear plasmid DNA. The resulting linear DNA was then ligated with the above fragment of about 7kb by a DNA ligase to give the plasmids pBOSMelα(+).

(3) Preparation of plasmid pBOSMelαδPro(+)

[0123] There was a Afill site in the boundary region between the precursor and metalloproteinase domains of Meltrin  $\alpha$ , and there was a Nhel site in the boundary region between metalloproteinase and disintegrin domains of Meltrin  $\alpha$ . On the other hand, there remined the Nhel site in the boundary region between the signal peptide-like sequence and disintegrin domain in pBOSMel $\alpha$ SMP(+) prepared in the above (1). Accordingly, pBOSMel $\alpha$  was digested at Afill, ligated with a Nhel linker immediately before its metalloproteinase domain and digested at Nhel, so that the metalloproteinase would be excised. The excised domain was inserted into the Nhel site between the signal peptide-like sequence and the disintegrin domain of pBOSMel $\alpha$ SMP(+) to give the expression plasmid, pBOSMel $\alpha$ SPro(+) encoding SPro wherein there a deletion was found around the precursor domain (the amino acid sequence of from Ile(No.55) to Glu(No.206) of Meltrin  $\alpha$ ).

(4) Confirmation of myoblast fusion-promoting activity

[0124] Myoblast cell line C2 was transfected by the mixture comprising the plasmid pBOSMelα(+) or pBOSMelαδMP (+), and the plasmid pSV2NEO in a molar ratio of 20:1 by using LIPOFECTAMINE (Gibco BRL) according to its protocol. The transfected cells were diluted and inoculated on a plate (φ 10cm) coated with collagen (IWAKI) so that the transformants would be obtained at a density of 10 - 20 clones per plate. The inoculated cells were cultured for 12 days in DMEM containing 20 % fetal bovine serum and 5 ng/ml of bFGF (Gibco BRL) followed by isolation thereof.

[0125] For the purpose of the examination of myoblast fusion-promoting activity, the resulting transformants and the parent strain C2 were cultured for 3-4 days in the absence of bFGF, inoculated onto a plastic dish (φ 6cm) at a density of 2 x 10<sup>5</sup>/dish, and further cultured for one day, followed by the 4 day culture in the above differentiation medium for differentiation induction. Upon differentiation induction, C2 began to form myotube. After the 4 day culture followed by fixation with methnol and staining with Giemsa and Wright's reagents (Merck), the number of nuclei were determined at any four independent fields of 1 mm<sup>2</sup> on the dish and fusion index was calculated as follows:

Fusion Index = 100 \* (The number of nuclei in multicleate syncytium having three or more nuclei) / (The number of the total nuclei)

[0126] Further, the time course of the fusion index was observed after differentiation induction every one day for five days.

[0127] The results are shown in Fig.11a - Fig.11b. As seen from these figures, the fusion activity of the transformant expressing the full length of Meltrin  $\alpha$  (pBOSMel $\alpha$ (+) which was referred to as "full length" in Fig.11a) become lower than that of the parent cell, and it was therefore considered that the full length of Meltrin  $\alpha$  would suppress the cell fusion in some way. On the other hand, the transformant harboring pBOSMel $\alpha$ 8MP(+), which was referred to as " $\Delta$ MP" in the figures, significantly promoted the cell fusion activity. It was also observed that the transformant harboring pBOSMel $\alpha$ 8Pro (+) promoted the cell fusion activity.

[0128] On the other hand, the C2 cell transformed by the plasmid pBOSMel $\beta$ (+) prepared by the insertion of the DNA encoding the full length of Meltrin  $\beta$  in the same way as in the above (2) could not cause any significant change in the fusion activity for muscle cells. However, The C2 transformant cotransfected by pBOSMel $\alpha$ (+) and pBOSMel $\beta$ (+) promoted the cell fusion activity compared with that of parent cell.

[0129] On the other hand, neither the C2 cell transformed by the plasmid pBOSMel $\gamma$ (+) prepared by the insertion of the DNA encoding the full length of Meltrin  $\gamma$  in the same way as in the above (2), nor the C2 transformant cotransfected by pBOSMel $\alpha$ (+) and pBOSMel $\gamma$ (+) could cause any significant change in the fusion activity for muscle cells.

[0130] These results demonstrate that Meltrin  $\alpha$  is involved in the fusion of muscle cells, and will show its activity to promote the cell fusion upon its processing. It is estimated that Meltrin  $\alpha$  or Meltrin  $\beta$  does not act alone, but act in the form of a heteromer between them, since the transformant expressing both Meltrin  $\alpha$  and Meltrin  $\beta$  promoted the fusion of muscle cells.

(5) Examination of the function of Meltrins in non-muscle cells

[0131] The mouse fibroblast L929 was transformed by pBOSMel $\alpha$ (+) or pBOSMel $\beta$ (+) and the transformants expressing Meltrin  $\alpha$  or Meltrin  $\beta$  were isolated. These transformants did not aggregate, nor fuse with each other. This was also true for the case of the transformant expressing both Meltrin  $\alpha$  and Meltrin  $\beta$ .

[0132] On the othe hand, the L929 cells transformed by pBOSMely(+) could showed a significant aggregation activity upon the addition of calcium ion, after the cells had been torn from a plate in a medium comprising no calcium ion.

[0133] These results demonstrate that Meltrin  $\gamma$  has a cell aggregation activity, and by considering the similarity of these molucules it is suggested that myoblast fusion-promoting activity of Meltrin  $\alpha$  and Meltrin  $\beta$  may be attributed to their myoblast aggregation-promoting activity.

Example 5: Inhibition of adhering activity by antisense

20

[0134] The plasmid BOSMelαδMP(-) prepared in Example 4 (1) was mixed with the plasmid PSV2NEO at a molar ratio of 20:1, by which C2 cells were tranformed according to the method of Example 4 (4) followed by isolation of the transformants expressing antisense RNA. The adhering activity of the thus isolated transformants was determined by the method of Example 4. The results are shown in Fig.11a - Fig.11b, which demonstrated that the fusion of C2 cells was inhibited by the expression of antisense RNA for δMP (referred to as "AS" in the figures).

[0135] The above results have revealed that Meltrin α plays an essential role in the cell fusion of muscle cells.

Example 6: Preparation of cDNA fragments encoding human Meltrins  $\alpha$  and  $\gamma$ 

[0136] By using mRNA purified from human myelocytes (Clonetech Co.) as a template, cDNAs were prepared according to the method of Example 1 (1), and 36 cycles of PCR was then carried out by using the degenerative primer obtained in Example 1 (2) and said cDNAs as a template. The amplified product was inserted into a EcoRV site of pBS-SKII(-), and named "pBShuMα300." The results of DNA sequencing are shown in Fig.12a and Fig.12b.

[0137] It was found that the DNA sequence comprised the base sequence encoding the part from an intermediate position of the disintegrin domain to an intermediate position of the cysteine-rich region of human Meltrin  $\alpha$  (the disintegrin domain is located to Gly (No.36), followed by the cysteine-rich region in Fig.12a and Fig.12b).

[0138] On the other hand, by using a part of a human sequence (D-14665) registered with a data base, whose function had not yet identified, a senseprimer (5'-CACGATGATGGGAGAGATTG-3') and antisense primer (3'-CACTCTGATT-TCCTATGCCTC-5') were synthesized. PCR was carried out according to the above method to give the amplified product, which was then inserted into the EcoRV site of pBS-SKII(-), and named "pBShuMyG238." The results of DNA sequencing are shown in Fig.13a and Fig.13b.

[0139] It was found that the DNA sequence comprised the base sequence encoding the part from an intermediate position of the metalloproteinase domain to an intermediate position of the cysteine-rich region of human Meltrin γ (the metalloproteinase domain is located from N-termial to Pro (No.40), the disintegrin domain from Lys (No.41) to Gly (No. 136) or from Tyr (No.46) to Gly (No.136), followed by the cysteine-rich region from Tyr (No.137)). The E. coli strain JM109 was transformed by those plasmids to give JM109(pBShuMα300) and JM109(pBShuMγG238), which were deposited in the National Institute of Bioscience and Human-Technology, Agency of Industrial Science and Technology

(1-3, Higashi 1-chome, Tsukuba-shi, Ibaraki-ken 350 Japan) on February 19, 1996 under accession numbers FERM P-15454 and 15455, respectively, and then transferred on October 8, 1996 to the deposit under the terms of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure and Regulation under accession numbers FERM BP-5704 and 5705, respectively.

Example 7: Preparation of cDNA fragment encoding human Meltrin  $\alpha$  by usig cDNA library derived from human placenta -1

(1) First screening

5

10

20

25

40

[0140] Based on the cDNA sequence of Meltrin  $\alpha$  obtained in Example 6, sense primer MA-1 and antisense primer MA-2 were synthersized (see Table 1). The human placenta  $\lambda$ gt11 cDNA library (Clonetech Co., code No. CLHL1008b) was inoculated onto LB plate ( $\phi$  10cm) at such a density that 10,000 plaques per plate may be obtained. After the formation of plaques, SM buffer 5ml was added to each plate, the plates were put by incubation at a room temperature for 4 hours, and phages were collected from each plate (plate lysate method). PCR was carried out by using the collected phage solution as a template. Thus, mA-1 and MA-2 primers, Ex Taq polymerase (TaKaRa Co.,), and its reagents (TaKaRa Co.,) were mixed, followed by 35 cycles of the reactions at 94°C for 30sec, 55°C for 30sec, and 72°C for one min. A part of the amplified products was subjected to an agarose gel electrophoresis, and a phage solution of the clone comprising Meltrin  $\alpha$  cDNA was selected.

(2) Second screening

[0141] The phage solution of the desired clone obtained in the first screening was inoculated at such a density that 400 plaques per plate may be obtained. After the formation of plaques, phages were collected in the same manner as above and a phage solution comprising the desired clone was selected.

- (3) Third screening
- [0142] The phage solution of the desired clone obtained in the second screening was inoculated at such a density
  that 40 plaques per plate may be obtained. After the formation of plaques, phages were collected in the same manner
  as above and a phage solution comprising the desired clone was selected.
  - (4) Forth screening
- [0143] The phage solution of the desired clone obtained in the third screening was inoculated at such a density that 10 plaques per plate may be obtained. After the formation of plaques, phages were collected in the same manner as above and a phage solution comprising the desired clone was selected.
  - (5) Final screening

[0144] The phage solution of the desired clone obtained in the forth screening was inoculated at such a density that 20 plaques per plate may be obtained. After the formation of plaques, each plaque was stuck with a toothpick, and the sticking material was suspended as a template into PCR solution. The above 35 cycles of the PCR with MA-1 and MA-2 primers finally gave two psitive clones. A single positive plaque comprising the desired clone was collected in SM buffer, and the phage was lysed thereinto.

[0145] PCR was carried out by using  $\lambda$ gt11 Forward primer and  $\lambda$ gt11 Reverse primer (Table 1) to give a fragment of human Meltrin  $\alpha$  cDNA in the phage vector.

[0146] From a partial DNA sequencing of the terminal bases of the resulting fragments it was estimated that those cDNAs comprised the base sequences encoding human Meltrin α obtained in Example 6, and corresponded to about 650 amino acids (Clone 23) or about 500 amino acids (Clone 25) of mouse Meltrin (Fig. 14).

Example 8: Preparation of cDNA fragment encoding human Meltrin α by usig cDNA library derived from human placenta -2

[0147] A sense primer Mel α-5'S was designed based on the sequence encoding the N-terminal of the cDNA sequence of the clone 23 revealed in Example 7. The human placenta λgt11 cDNA library (Clonetech Co.) was screened by the sense primer Mel α-5'S and antisense primer MA-2 to give cDNA encoding about 700 amino acids (Clone 26) (Fig.14a). For the purpose of the analysis of the base sequence of Meltrin gene, the four primers, λgt11 Forward-Eco, λgt11

Reverse-Eco, MA-1-Eco, and MA-2-Eco were synthesized (Table 1).

5

10

15

50

[TABLE 1] The base sequences of the primers for PCR

MA-1	: 5' ACG ATG GGC ACT CAT GTC AG 3'
MA-2	: 5' CAT CTC GCA TTT GGC AAA GG 3'
λ gt11 Forward	: 5' GGT GGC GAC GAC TCC TGG AGC CCG 3'
λ gt11 Reverse	: 5' TTG ACA CCA GAC CAA CTG GTA ATG 3'
Mel α-5' <i>S</i>	: 5' CAC TGA ACA TTC GGA TCG TG 3'
λ gt11 Forward-Eco	:5'CCGGAATTCGGTGGCGACGACTCCTGGAGCCCG3
λ gt11 Reverse-Eco	: 5' CCG GAA TTC TTG ACA CCA GAC CAA CTG GTA ATG 3
MA-1-Eco	: 5' CCG GAA TTC ACG ATG GGC ACT CAT GTC AG 3'
MA-2-Eco	: 5' CCG GAA TTC CAT CTC GCA TTT GGC AAA GG 3'
S-hMel α-TM5'	: 5' GCA CAA AGT GTG CAG ATG GA
A-mMel α-3'	: 5' CAG AGG CTT CTG AGG AGG N

[0148] The second half of the Meltrin gene was amplified by PCR using Clone 25 as a template, and MA-1-Eco and  $\lambda$ gt11 Reverse-Eco primers. The first half of the Meltrin gene was amplified by PCR using Clone 26 as a template, and MA-2-Eco and  $\lambda$ gt11 Forward-Eco primers. These cDNA fragments were digested at EcoRI and cloned into the EcoRI site of pUC 118 to give the plasmid vectors "pMel $\alpha$ -26N" and "pMel $\alpha$ -25C", respectively. The sequences of Meltrin  $\alpha$  cDNA comprised in these plasmids were determined by a conventional method.

[0149] The E.coli strain JM109 was transformed by those plasmids according to the known method of Hanahan et al. to give JM109(pMel $\alpha$ -26N) and JM109 (pMel $\alpha$ -25C), and were deposited in the National Institute of Bioscience and Human-Technology, Agency of Industrial Science and Technology (1-3, Higashi 1-chome, Tsukuba-shi, Ibaraki-ken 350 Japan) on October 3, 1996 under the terms of the Budapest Treaty on the International Recognition of the Deposit of microorganisms for the Purposes of Patent Procedure and Regulation under accession numbers FERM BP-5689 and 5688, respectively.

[0150] The base sequence and its corresponding amino acid sequence of human Meltrin  $\alpha$  which had been revealed by the base sequencing of pMel $\alpha$ -26N and pMel $\alpha$ -25C are shown in Fig.15a - Fig.15f.

[0151] Comparison of the DNA sequence thus obtained with that obtained in Example 6 indicated four discrepancies in base pairs, the three of which being silent mutation, and the other dicrepancy causing substitution of Asp (No.505) in the above figures for Glu in the sequence of Example 6.

[0152] The analysis of the structure of the base sequence showed that the DNA encoded the sequence from an intermediate part of the precursor domain to the C-termial of Meltrin  $\alpha$ . Thus, it has been considered that in the amino acid sequence shown in Fig.15a - Fig.15f, the partial sequence (C-terminal) of the precursor domain corresponds to the the sequence from Gly N-terminal to Arg (No.155) and to the bases No.1-465, the metalloproteinase domain to the sequence from Glu (No.156) to Pro (No.364) and to the bases No.466-1092, the disintegrin domain to the sequence from Glu (No.365) or Phe (No.370) to Gly (No.459) and to the bases No.1093 or 1108-1377, the cysteine-rich region to the sequence from His (No.460) to Gln (No.656) or Ala (No.652) and to the bases No.1378-1968 or 1956, the fusion peptide-like sequence to the sequence from Gly (No.535) to Gln (No.557) and to the bases No.1603-1671. There was no transmembrane domain in this sequence, suggesting that human Meltrin  $\alpha$  existed as a soluble protein without a transmembrane domain in a body. In other words, it is considered that Meltrin  $\alpha$  having the amino acid sequence of Fig. 15a - Fig.15f is extracellularly secreted and present in blood or body fluid. It is considered that such soluble Meltrin  $\alpha$  takes a part in regulating adhesion, fusion and aggregation of cells in the body.

[0153] It is considered that Meltrin  $\alpha$  having the amino acid sequence of Fig.15a - Fig.15f has generated as a result of an alternative splicing of the gene. It is also considered that the DNA encoding the region downstream of the cysteinerich region, and the DNA encoding transmembrane doamin and intracellular domain are located on different exons, and that the splicing out of either DNA would yield a soluble type Meltrin, or a membrane-binding type Meltrin.

Example 9: preparation of cDNA fragments encoding human Meltrins β

- (1) Preparation of cDNA fragment encoding a part of the disintegrin domain of human Meltrin  $\beta$
- 65 [0154] By using mRNA purified from human myelocytes (Clonetech Co.) as a template, cDNAs were prepared according to the method of Example 1 (1), and 36 cycles of PCR were then carried out by using the degenerative primers obtained in Example 1 (2) and said cDNAs as a template. The amplified product was inserted into pBS-SKII(-). The analysis of the resulting DNA sequence revealed that it was a partial sequence of Meltrin β. The determined DNA

sequence is shown in Fig.16.

- (2) First screening by using cDNA library originated in human fetal lung
- [0155] Based on the partial cDNA sequence of Meltrin β obtained in the above (1), sense primer MA-3 and antisense primer MA-4 were synthersized (see Table 2). The human fetal lung λgt11 cDNA library (Clonetech Co., code No. CLHL1072) was inoculated onto LB plate (φ 10cm) at such a density that 10,000 plaques per plate may be obtained. After the formation of plaques, SM buffer 5ml was added to each plate. And the plates were put at a room temperature for 4 hours, and phages were collected from each plate (plate lysate method). PCR was carried out by using the collected phage solution as a template. Thus, MA-3 and MA-4 primers, Ex Taq polymerase (TaKaRa Co.,), and its reagents (TaKaRa Co.,) were mixed, followed by 35 cycles of the reactions at 94°C for 30sec, 55°C for 30sec, and 72°C for one min by means of DNA thermal cycler (Perkin Elmer Co.,). A part of the amplified products was subjected to an agarose gel electrophoresis, and a phage solution of the clone comprising Meltrin β cDNA was selected.
- 15 (3) Second screening
  - [0156] The phage solution of the desired clone obtained in the first screening was inoculated at such a density that 1000 plaques per plate may be obtained. After the formation of plaques, phages were collected in the same manner as above and a phage solution comprising the desired clone was selected.
  - (4) Third screening

20

25

- [0157] The phage solution of the desired clone obtained in the second screening was inoculated at such a density that 100 plaques per plate may be obtained. After the formation of plaques, phages were collected in the same manner as above and a phage solution comprising the desired clone was selected.
  - (5) Forth screening
- [0158] The phage solution of the desired clone obtained in the third screening was inoculated at such a density that
  10 plaques per plate may be obtained. After the formation of plaques, phages were collected in the same manner as
  above and a phage solution comprising the desired clone was selected.
  - (6) Collection and confirmation of DNA fragment comprising partial cDNA sequence
- 35 [0159] The PCR was carried out using the phage solution of the desired clone obtained in the forth screening (#24) as a template, and a combination of λgt11 Forward primer (Table 1) and MA-4 primer or a combination of λgt11 Reverse primer (Table 1) and MA-3 primer to give amplified products with about 500bp (24-F/4) and about 5kbp (24-R/3), respectively. From a partial DNA sequencing of the terminal bases of the resulting two DNA fragments, it was estimated that those cDNA comprised the base sequences determined in the above (1).
  - (7) Analysis of base sequences
  - [0160] For the purpose of subcloning of the cDNA fragments comprising the cDNA partial sequence of human Meltrin β, two primers MA-3-Eco and MA-4-Eco were newly synthersized (see Table 2).
- [0161] The PCR was carried out using the phage solution (#24) as a template, and a combination of λgt11 Forward-Eco primer (Table 1) and MA-4-Eco primer or a combination of λgt11 Reverse-Eco primer (Table 1) and MA-3-Eco primer. The resulting amplified products were digested with EcoRI and inserted into the EcoRI site of pUC118 to give the plasmids, "pMeIβ-24C" and "pMeIβ-24N", respectively. The sequence of Meltrin β cDNA comprised in these plasmids was determined by a conventional method.
- 50 [0162] The E. coli strain JM109 was transformed by those plasmids according to the known method of Hanahan et al. to give JM109(pMelβ-24C) and JM109 (pMelβ-24N), and were deposited in the National Institute of Bioscience and Human-Technology, Agency of Industrial Science and Technology (1-3, Higashi 1-chome, Tsukuba-shi, Ibaraki-ken 350 Japan) on October 3, 1996 under the terms of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure and Regulation under accession numbers FERM BP-5690 and 5691, respectively.
  - [0163] The base sequence and its corresponding amino acid sequence which had been revealed by the base sequencing of pMelβ-24C and pMelβ-24N are shown in Fig.24a Fig.24e.
  - [0164] Comparison of the DNA sequence thus obtained with that obtained in the above (1) showed one discrepancy

in base pairs, which was a silent mutation, causing no change of amino acid.

[0165] The analysis of the structure of the base sequence showed that the DNA encoded the sequence from an intermediate part of the metalloproteinase domain to the C-termial of human Meltrin  $\beta$ . Thus, it has been considered that in the sequence shown in Fig.24a - Fig.24e, the partial sequence at C-terminal of the metalloproteinase domain corresponds to the the sequence from Gly (N-terminal) to Pro (No.36) and to the bases No.2-109, the disintegrin domain to the sequence from Asp (No.37) or Tyr (No.42) to Gly (No.131) and to the bases No.110 or 125-394, the cysteine-rich region to the sequence from Thr (No.132) to Pro (No.330) and to the bases No.395-991, the transmembrane domain to the sequence from Val (No.331) to Met (No.348) or Arg (No.353) and to the bases No.992-1045 or 1060. It is considered that the sequence from Tyr (No.349) or Gln (No.354) corresponds to the intracellular domain. However, as homology analysis to mouse Meltrin  $\beta$  shows a very low homology in the sequence from Pro (No.395), it is estimated that the sequence up to His (No.394) is involved in the function of extracellular domain of human Meltrin  $\beta$ . The sequence up to Pro (No.395) in Fig. 24a - Fig.24e is shown in Fig.17a - Fig.17c.

[TABLE 2] The base sequences of the primers for PCR

[	SEE 2) The base coductions of the printer for the training
MA-3	: 5' TGC TGC CAC CAG TGT AAG 3'
MA-4	: 5' TCC TGG TAG GTG AGG CAC ATG 3'
MA-3-Eco	: 5' CCG GAA TTC TGC TGC CAC CAG TGT AAG 3'
MA-4-Eco	:5'CCG GAA TTC TCC TGG TAG GTG AGG CAC ATG 3'

Example 10: Preparation of anti-Meltrin  $\alpha$  monoclonal antibodies

## (1) Selection of peptides

15

20

25

45

[0166] Based on the amino acid sequence of mouse Meltrin  $\alpha$  determined in Exmple 1, their epitopes were analysed. [0167] Eight kinds of peptide sequences were selected as a potential epitope, based on the secondary structure estimated from the regions wherein discrepancy in amino acids is seen between Meltrins  $\alpha$  and  $\beta$ , the estimated non-RGD region, and the region wherein metalloproteinase had been cleaved (Fig.18a and b). These eight kinds of peptides were synthesized by Peptide Synthesizer (ABI 432A) so that they would have Cys at their C-terminal, cleaved, and purified by HPLC of a reverse phase column (YMC-ODS).

#### (2) Preparation of antiserum

[0168] After lyophilization of the peptides obtained in the above (1), each peptide 0.55mg was dissolved in 0.1 M phosphate buffer (pH 7.0) 55μl. Maleimidated KLH (Boehringer Manheim) 0.77mg was dissolved in distilled water 77μl. The two resulting solutions were combined, and reacted at a room temperature for two hours, followed by the purification by Nick column (Pharmacia) equilibrated with physiological saline to give antigens to be used in the following experiments.
 [0169] Each antigen 50μg was diluted with physiological saline to 0.1 ml, mixed with the same amount of Freund's complete adjuvant (DIFCO) and administered intraperitoneally into Wistar rat (5 weeks old, female). The antigen was mixed with the same amount of Freund's incomplete adjuvant (DIFCO) and administered two weeks later in the same way as above.

### (3) Evaluation of antiserum (plate assay)

[0170] After one week from the administration, the blood was drawn from the eyeground of the rat, and an increase of the antibody titer for the administered peptides was confirmed by the reaction between immobilized peptides and the antiserum according to a plate assay as follows.

[0171] First, 50mM phosphate buffered saline (0.9% NaCl, pH 7.2) comprising 0.5mg/ml of Sulfo-SMCC (Pierce) was poured into each well of an amino plate (Sumitomo Bakelite). After incubation at 37°C for 2 hours, the wells were washed five times with ion-exchanged water, and the above buffer comprising 0.5μg/ml of each peptide was added. After incubation at 37°C for one hour, the well were blocked by 0.076M phosphate buffered saline (0.45% NaCl, pH 6.4), which will be referred to hereinafter as "PBS", comprising 0.1% of BSA and 4mg/ml of cysteamine. The blocking agent was removed, each antiserum diluted 1,000 to 100,000 times with PBS was added followed by incubation at 37°C for one hour. After two repeats of washing of the wells with 0.9% NaCl comprising 0.005% Tween20, an anti-rat immunoglobulin abtibody labelled with peroxidase (Dako) and diluted with PBS comprising 10% rabbit serum was added to each well followed by incubation at 37°C for one hour. Upon the completion of the reaction, the wells were washed five times with

a washing liquid and two times with ion-exchanged water. And 0.1M McIlvaine buffer (pH 5.0) comprising 3mg/ml of ophenylene diamine and 0.027% hydro peroxide was added and reacted for 5min. The reaction was terminated by the addition of 1N HCl, and absorbance at 490nm was measured. The results are shown in Table 3, in which (++) means a strong reactivity, and (+) means a week reactivity.

TABLE 3 Reaction of antiserum with the peptide antigens

	_ · · ·
peptide antigens	Reaction of Antiserum
1 ProA	++
2 MP-A	++
3 MP-B	++ .
4 DC-A	· +
5 DC-B	+
6 DC-C	++
7 DC-D	N.D.
8 DEA	++
N.D. (not determined)	

(4) Evaluation of antiserum (Western blotting)

[0172] For the confirmation of the binding of the antiserum prepared in the above (2) to Meltrins, Western blotting was carried out.

[0173] Mouse myoblast C2 was transformed by pBOSMel $\alpha\delta$ Pro(+) and pBOSMel $\beta$ (+), which will be referred to hereinafter as "#9-3", and mouse myoblast C2 was transformed by pBOSMel $\alpha\delta$ MP(+), which will be referred to hereinafter as "#3-5."

[0174] The transformed C2 cells of 1 x 10<sup>7</sup> cells were washed with PBS-(GIBCO BRL) and collected by centrifugation. The density of the collected cells was adjusted to 5 x 10<sup>6</sup> cells/ml, mixed with a proteolysis inhibitor, C¢ mplete (Boehringer Manheim) in amount of one 25th of the volume of the cell-mixture, and mixed with SDS to a final concentation of 0.2%. After incubation at a room temperature for 30min, the cells were subjected to sonication at 4°C for 10sec (Isec x 10), and centrifuged. The resulting supernatant was collected and used as a cell lysate. Another cell lysate was prepared from fibroblast L929 (ATCC No.CCL-1) in the same way, and used as a negative control.

[0175] The resulting cell lysate (10µl) was mixed with an equiamount of a gel loading buffer (0.25M Tris-HCl, 2% SDS, 30% Glycerol, 0.01% BPB(pH 6.8)), the resulting solution (6µl) was applied to SDS-PAGE of 4-20T % (Tefco), and electrophoresed under 25mA at a room temperature for about one hour. After the completion of the electrophoresis, the contents were transferred to PVDF membrane (Millipore) under the conditions of 150mA, 4°C and 45min. The membrane was blocked by shaking in 4% skim milk (Meiji Milk Co.) at a room temperature for one hour, and each lane was cut. Each excised lane was soaked and shaked in antiserum (1ml) diluted 500 times with 50mM Tris-HCl (pH 7.2) comprising 0.05% Tween20 (referred to hereinafter as "T-TBS") and 4% skim milk at a room temperature for one hour. After the completion of the reaction, each lane was washed two times with T-PBS, soaked in 1ml of an anti-rat immunoglobulins antibody labelled with HRPO (Dako) diluted 500 times with T-PBS comprising 4% skim milk, and reacted at a room temperature for one hour. After washing five times with T-PBS, it was detected by ECL system (Amersham). The results are shown in Table 4. Bands were detected in the three kinds of the antiserums by the Western blotting.

TABLE 4 Reaction of antiserum with the cell lysate in Western blotting

Peptide antigens	Western blottting					
1 ProA	+					
2 MP-A	-					
3 MP-B	-					
4 DC-A	N.D.					
5 DC-B	N.D.					
6 DC-C	+					
7 DC-D	N.D.					

50

5

10

15

20

35

40

45

(continued)

Peptide antigens	We	stern blottting
8 DEA	+	
N.D. (not determined)	- · - · ·	

### (5) Preparation of monoclonal antibody

5

20

25

[0176] The antigens (ProA, MP-B, DC-C, DEA) (50µg each) were diluted with 400µl of physiological saline, and injected into the tail vein of the rats whose antibody titer had increased. Three days later, cell fusion was carried out by using myeloma P3X63Ag8U.1 according to the known method (Monoclonal antibody Jikken Sosa Nyumon (Guide of monoclonal antibody preparation), Tamie Ando and Jo Chiba, Koudan-sha Scientific). Six days later, the culture supernatant was collected and subjected to the plate assay according to the method of the above (3). The wells that showed reactivity with the peptide antigens were subjected to cloning by limiting dilution(Monoclonal antibody Jikken Sosa Nyumon (Guide of monoclonal antibody preparation), Tamie Ando and Jo Chiba, Koudan-sha Scientific). After cloning, the screening by the plate assay was performed again to give 27 clones of the hybridomas producing an anti-mouse Meltrin α monoclonal antibody which reacted with the peptide antigens. The results are shown in Table 5.

TABLE 5 Hybridomas producing anti-Meltrin peptides monoclonal antibody

Peptide antigens	Hybridoma No.	The number of Hb
ProA	F936	10
MP-B	F939	4
DC-C	F933	4
DEA	F934	. 8

[0177] Purified antibodies were obtained from the thus established anti-Meltrin monoclonal antibody-producing hybridoma cell lines by the following method.

[0178] The hybridomas were cultured in RPMI1640 supplemented with 10% fetal bovine serum and 1ng/ml of human IL6 till a final density of 2 x 10<sup>5</sup> cells/ml. The medium was then exchanged with a serum-free medium (Hybridoma-SFM, GIBCO BRL), and the culture was continued until the cells died. The resulting culture supernatant was filtered through filter paper for the removal of the cells, and subjected to purification by Protein G column (Prosep-G, Bioprocessing INC) as follows. The culture supernatant (1L) was applied to Prosep-G column (20ml) at a flow rate of 10ml/min, followed by washing with 0.1M phosphate buffer (pH 7.5) comprising 0.15M NaCl. After the absorbance at 280nm had decreased, the bound monoclonal antibody was eluted by 0.1M citric acid buffer (pH 3.0). After neutralization of the pH, the eluate was concentrated with DIAFLO (Grace Japan), and dialysed against 0.076M phosphate buffered saline (pH 6.4) comprising 0.45% NaCl. The concentration of the purified antibody was calculated on the basis of the absorbance at 280nm.

### (6) Evaluation of monoconal antibody

[0179] The binding activity of 7 lots of the purified antibodies (10 µg/ml each) obtained in the above (5) to Meltrin was confirmed by Western blotting according to the method of the above (4) using the cell lysate of #9-3 cell. The results are shown in Fig. 19. The band of about 67kDa specific to the cell lysate of #9-3 cell was detected by the reaction with F933-4-3 (subclass IgG2a), F933-10-26 (subclass IgG2a), F934-17-6 (subclass IgG2a), F934-3-23 (subclass IgG2a), F934-4-33 (subclass IgG2a), F934-6-3 (subclass IgG2a), and F934-20-5 (subclass IgG2c). As these bands were not detected in the case of the cell lysate of L929 cell, it was confirmed that the monoclonal antibodies obtained in the above (5) were bound to Meltrin.

## Example 11: Preparation of anti-mouse Meltrin monoclonal antibody

(1) Preparation of the antigen to be adminitered and immunization of rat

[0180] Rats were immunized with #9-3 and #3-5 cells as the antigen to be administered as follows. The cells used as the antigen to be administered were cultured in the absence of bFGF. First, the cells cultured in four dishes to a density of about 5 x 10<sup>5</sup> cells /\phi 10cm dish were subcultured in 20 dishes to until the same density as the above, then again subcultured in 40 dishes (\phi 15cm) up to a density of about 5 - 6 x 10<sup>6</sup> cells / dish, and futher cultured in a differentiation

medium (DMEM supplemented with 2% horse serum) for two days to finally form myotube. These cells were then scraped with a silicon rubber Policeman, washed two times with PBS, and suspended into the medium comprising 10% DMSO for storage at -80°C.

[0181] The #9-3 and #3-5 cells were suspended in physiological saline (200 $\mu$ l), mixed with an equiamount of Freund's complete adjuvant (DIFCO) and intraperitoneally administered into Wistar rat (5 weeks old, female) in an amount of 1 x 10<sup>7</sup> cells/rat. The antigen was mixed with the same amount of Freund's incomplete adjuvant (DIFCO) and administered two weeks later in the same way as above.

#### (2) Evaluation of antiserum

10

30

35

40

50

[0182] After one week from the boosting, the blood was drawn from the eyeground of the rat, and a binding of antiserum to Meltrin was determined by using the cell extract according to the plate assay of Example 10 (3). The cell extracts of #9-3, #3-5 and L929 cells were prepared according to the method of Example 10 (4), except that NP-40 (Nacarai Tesque Co.) was used at a final concentration of 0.5% as a surfactant.

[0183] First, each cell extract was dilulted with PBS to a concentration of 40 μg/ml, each 50μl of which was separately poured into each well of an immuno plate (Maxisorp Nunc). After incubation at 56°C for 30min for binding of the antigen, the wells were washed five times with ion-exchanged water, blocked by 20 % Block Ace (Yukijirushi Milk Co.) / PBS 100μl, followed by incubation at a room temperature for 30min. After removal of the blocking agent, each antiserum (50μl) was added and incubated at 37°C for one hour. After two repeats of washing of the wells with the washing liquid, 50μl of an anti-rat immunoglobulins antibody labelled with peroxidase (Dako) and diluted 1,000 times with 10% Block Ace /PBS was added to each well followed by incubation at 37°C for one hour. Upon the completion of the reaction, the wells were washed five times with the washing liquid and two times with ion-exchanged water, and 50μl of 0.1M McIlvaine buffer (pH 5.0) comprising 3mg/ml of o-phenylene diamine and 0.027% hydro peroxide was added and reacted for 10 min. The reaction was terminated by the addition of 1N HCl (50 μl), and the absorbance at 490 nm was measured.

[0184] Western blotting was also carried out by using the cell extract of L4-3 described in the following (4) to confirm its binding to Meltrin. The results are shown in Table 6.

[0185] It was confirmed that the antiserum obtained from the rats immunized with #9-3 and #3-5 cells reacted with the corresponding cell extract, and were bound to Meltrin in the Western blotting.

TABLE 6 Reaction of antiserum of the rats immunized with #9-3 and #3-5 cells to Meltrin

L4-3			
+			
+			
-			

## (3) Preparation of monoclonal antibody

[0186] The #9-3 and #3-5 cells (  $1 \times 10^7$  cells each) were suspended in physiological saline (200 $\mu$ I), and intraperitoneally administered into the rat whose antibody titer had increased. Three days later, cell fusion was carried out by using myeloma P3X63Ag8U.1 according to the known method (Monoclonal antibody Jikken Sosa Nyumon (Guide of monoclonal antibody preparation), Tamie Ando and Jo Chiba, Koudan-sha Scientific). Six days later, the culture supernatant was screened by its reactivity with the immobilized cell extracts. The wells that showed reactivity with the cell extracts were subjected to cloning by limiting dilution (Monoclonal antibody Jikken Sosa Nyumon (Guide of monoclonal antibody preparation), Tamie Ando and Jo Chiba, Koudan-sha Scientific). After cloning, the above screening was repeated to give 13 clones, 5 clone from the rat immunized with #9-3 ( $\delta$ Pro; hybridoma No. F932) and 8 clones from the rat immunized with #3-5 ( $\delta$ MP; hybridoma No. F937).

# (4) Evaluation of monoconal antibody

[0187] The monoclonal antibodies F932-15-2 (subclass IgG1) and F937-9-2 (subclass IgG1) that showed a high reactivity with the cell extracts were evaluated.

[0188] First, the staining of myotube formed by C2 cells was examined by a cell immunofluorescence staining method. C2 cells were suspended in 10% FCS/DMEM at a density of 3 x 10<sup>4</sup> cells/ml, each 100µl of which was then separately poured into the wells of chamber slide (Lab-TEK, Nunc Co.). After the culture at 37°C and 5% CO<sub>2</sub> for two days, the medium was exchanged with 2% horse serum/DMEM. The cell staining was carried out by using myotube formed two

days later. The cells were washed two times with PBS<sup>-</sup>, and 4% formaldehyde was added followed by the reaction at a room temperature for 30min to fix the cells. The cells were washed three times with PBS<sup>-</sup> and blocked with 20% Block Ace /T-PBS. After removal of the blocking agent, antibodies diluted to 10µg/ml with 20% Block Ace /T-PBS was added and reacted at a room temperature for one hour. After three repeats of washing of the wells with PBS<sup>-</sup>, an anti-rat immunoglobulins antibody FITC (Dako) diluted 20 times with 10% rabbit serum/T-PBS was added to each well followed by incubation a room temperature for one hour. After the completion of the incubation, the cells were washed three times with PBS<sup>-</sup>, and subjected to fluorescence microscopy. It was observed that myotube was stained by both the antibodies, but not stained by rat IgG (ZYMED) used as a negative control.

[0189] Next, L929 cells experssing mouse Meltrin  $\alpha$  or  $\beta$  were prepared and subjected to cell staining for the purpose of confirmation of the specificity of the above antibodies. Thus, fibroblast L929 was transfected with the mixture comprising the plasmids pBOSMel $\alpha$ (+) and pBOSMel $\beta$ (+) prepared in Example 4, and the plasmid pSV2NEO in a molar ratio of 12:12:1 by using LIPOFECTAMINE (Gibco BRL) according to its protocol to give L4-3 cells expressing mouse Meltrins  $\alpha$  and  $\beta$ . Similarly, 1929 was transfected with the mixture comprising the plasmids pBOSMel $\beta$ (+) and the plasmid pSV2NEO in a molar ratio of 20:1 to give L2-10 cells expressing mouse Meltrin  $\beta$ . Similarly, L929 was transfected with the plasmids pBOSMel $\alpha$ Pro(+) to give L8-5 cells expressing mouse Meltrin  $\alpha$ Pro. The transfected cells were cultured in 10% FCS/DMEM and subcultured onto a chamber slide. The specificity of the antibodies was confirmed by cell staining using L929, L4-3, L2-10 and L8-5 cells. The results shown in Table 7 indicated that F932-15-2 was bound to Meltrins  $\alpha$  and  $\beta$ , and F937-9-2 was bound to Meltrin  $\alpha$ .

[0190] The hybridoma expressing the monoclonal antibody F932-15-2 was deposited with the National Institute of Bioscience and Human-Technology, Agency of Industrial Science and Technology (1-3, Higashi 1-chome, Tsukuba-shi, Ibaraki-ken 350 Japan) on October 3, 1996 under the terms of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure and Regulation under accession numbers FERM BP-5687.

T A			_
	ж	-	•

Cell	Expression	F932-15-2	F937-9-2
L929	•	•	-
L4-3	$\alpha$ and $\beta$	+	+
L2-10	β	+	-
L8-5	α (δΡιο)	+	+

(5) Determination of neutralizing activity

25

30

35

40

[0191] The neutralizing activity of the monoclonal antibodies obtained in the above (3) was confirmed by their inhibition of the formation of myotube by C2 cells. C2 cells were cultured in a collagen-coated dish contianing 10% FCS/DMEM till 80% of confluence, followed by exchange of the medium with 2% horse serum/DMEM supplemented with 0 or  $40\mu g/ml$  of the antibodies to be tested. The formation of byotube was then observed and the ratio of nuclei in the formed myotube was calculated. As seen from Fig.20, the formation of myotube on the day 2 was inhibited, showing that both F932-15-2 and F937-9-2 have the neutralizing activity.

Example 12: The activity of Meltrin neutralizing antibodies to inhibit the formation of bone resorption area (pit) in mouse unfractionated bone cells

[0192] Femur and tibia extracted from 13-day-old ICR mouse were crushed in MEM  $\alpha$  medium (GIBCO) supplemented with 5% FBS. After being allowed to stand still for 2min, the precipitaed bone residues were removed. The supernatant of the suspending cells was adjusted to 1 x  $10^7$  cells/ml,  $100\mu$ l of which was then added to each well of a 96 well microplate provided with ivory fragments. The ivory fragments had been thinly sliced, punched into 6mm in diameter, washed with 70% ethanol and sterilized. The mouse Meltrin-neutralizing antibody (F932-15-2) obtained in Example 11, and rat IgG were diluted with MEM  $\alpha$  medium (GIBCO) supplemented with 5 % FBS to final concentrations of 5, 50, and  $500\mu$ g/ml,  $100\mu$ l of which was then added to each well. After incubation at 37°C and 5% CO $_2$  for three days, the cells were removed with a scraper, and resorption area was stained with an acid hematoxylin solution (SIGMA) for about 7min and the number of the stained resorption area was caluculated using an ocular micrometer under a microscope by counting the number of squares wherein resorption fossa was contained.

[0193] The results are shown in Fig. 21, which demonstrates that the number of the formed resorption area was inhibited in a dose-depending manner by the mouse Meltrin-neutralizing antibody. Accordingly, it was suggested that the Meltrin-neutralizing antibody would affect directly or indirectly osteoclast and inhibit bone resorption.

#### Example 13: Serum Ca-decreasing activity of Meltrin-neutralizing antibody in mouse having enhanced bone resorption

[0194] Seven-week-old ICR mice (male) were fed for five days with low Ca feed with Ca content of 0.02% or less. The mouse Meltrin-neutralizing antibody (F932-15-2) obtained in Example 11 was injected into the tail vein of the mice (one group consisting of five mice) at doses of 0.1mg and 1mg per mouse). Rat IgG (Img per mouse) and phosphate buffer physiological saline were also injected as a control in the same way. Before injection and one day later, the blood was collected from the vein under eyes, and serum was separated. The value of Ca in the serum was then determined by an autoanalyzer (COBAS FARAII, ROCHE) using Ca determination kit (CalciumHR-II, WAKO Pure Pharmaceuticals). The results are shown in Fig.22.

[0195] As seen from Fig.22, the serum Ca value after one day from the injection in the groups treated with the mouse Meltrin-neutralizing antibody was lower than that of the groups treated with rat IgG or physiological saline. These results suggested that the Meltrin-neutralizing antibody would inhibit an unhealthly enhanced bone resorption due to hyperparathyroidism or malignant hypercalcemia.

## 5 Example 14: Preparation of cDNA fragment encoding human Meltrin α comprising transmembrane domain

[0196] A sense primer S-hMel $\alpha$ -TM5'was synthesized based on the partial cDNA sequence of human Meltrin  $\alpha$  obtained in Example 8, and an antisense primer A-mMel $\alpha$ -3' was synthesized based on the cDNA sequence of mouse Meltrin  $\alpha$  (see Table 1).

20 [0197] PCR was carried out by mixing the human placenta λgt11 cDNA library (Clonetech Co., code No. CLHL1008b) as a template, with S-hMelα-TM5' and A-mMelα-3' primers, Ex Taq polymerase (TaKaRa Co.,), and its reagents (TaKaRa Co.,), followed by 35 cycles of the reactions at 94°C for 30sec, 55°C for 30sec, and 72°C for one min. The base sequencing of the resulting amplified fragment (clone TM) suggested that the fragment was a human cDNA fragment corresponding to about 220 amino acids comprising the transmembrane domain of mouse Meltrin.

5 [0198] The obtained base sequence and its corresponding amino acid sequence are shown in Fig.23a - Fig.23b.

#### Example 15: Acute toxicity test

35

[0199] The mouse Meltrin-neutralizing antibody (F932-15-2) obtained in Example 11 was injected into seven-week-old ICR male mice (one group consisting of five mice) at doses of 1mg and 3mg per mouse). Phosphate buffer physiological saline was also injected into a control group in the same way. Neither significant decrease of body weight nor side effect was observed in any group after the injection. No dead mouse was observed, either.

## Reference Example 1: Preparation of monoclonal antibody recognizing human Meltrin

(1) Preparation of antibody using a peptide having the amino acid sequence derived from human Meltrin as an antigen

[0200] In consideration of the results obtained in Example 10, the sequence "GKVSKSSFAKCEMRDAKC" corresponding to DC-C in the amino acid sequence of human Meltrin  $\alpha$  obtained in Example 8 was synthesized in the same way as in Example 10 (1), purified and conjugated with maleimidated KLH to give an antigen to be administered.  $20\mu$ g of the antigen was dissolved in 0.1ml of physiological saline and mixed with an equiamount of FCA followed by injection to ddy mouse (5 weeks old, female). The same amount of the antigen was mixed with FIA and injected two weeks later. The blood was collected from the eyeground one week later and antiserum was prepared. Evaluation of the reactivity of the resulting antiserum with the administered peptide according to the method of Example 10 (3) revealed its specific reactivity with the administered peptide. Accordingly, mouse, rat, hamster and the like are immunized with the peptide antigen, and monoclonal antibody may be prepared in the same manner as in Example 10 (5). Such antibody may also be used in Western blotting.

[0201] As it is estimated that the amino acid sequence in Fig.15a - Fig.15f is Meltrin  $\alpha$  of a soluble type, an antibody, which may be effectively used in the determination of soluble Meltrin in the body, may be prepared by immunization of a peptide having the amino acid sequence adjacent to C-termial of the above sequence.

[0202] Similarly, antibodies recognizing human Meltrin  $\beta$  and Meltrin  $\gamma$  may be prepared by chemically synthesizing peptides having the amino acid sequences of suitable parts in the amino acid sequences in Fig.17a - Fig.17c or Fig.13a - Fig.13d and injecting the thus synthesized peptides into animals. In any case, the amino acid sequence will be selected from the extracellular domain.

[0203] For the preparation of an antibody specific to each one of Meltrins  $\alpha$ ,  $\beta$  and  $\gamma$ , the amino acid sequence should be selected from the parts with a low homology among them, and a peptide having the thus selected amino acid sequence is synthesized and injected to animals such as mouse, rat and hamster in the same way as in Example 10 (2).

[0204] In any case, monoclonal antibodies are prepared in the same way as in Example 10 (5).

(2) Preparation of anti-Meltrin monoclonal antibody using cells expressing human Meltrin as an antigen

[0205] DNA encoding the amino acid sequence wherein the amino acid sequence located downstream of the transmembrane domain shown in Fig.23a - Fig.23b is fused downstream of the sequence from the metalloproteinase or the disintegrin domain to the cysteine-rich region shown in Fig. 15a- Fig. 15f is prepared, and inserted into an expression vector pEFBOS, followed by transformation of C2 cells by the resulting vector. The transformant is treated as in Example 11 (1), and used as an antigen for immunization of animals such as mouse, rat and hamster. Antibodies recognizing human Meltrin  $\alpha$  is screened as in Example 11 (2), and monoclonal antibodies are prepared as in Example 11 (3).

[0206] Similarly, DNA encoding the amino acid sequence shown in Fig.17a - Fig-17c or the sequence located downstream of the disintegrin domain of the above sequence is prepared, and inserted into an expression vector pEFBOS, followed by transformation of C2 cells by the resulting vector. The transformant is treated as in Example 11 (1), and used as an antigen for immunization of animals such as mouse, rat and hamster. Antibodies recognizing human Meltrin β is screened as in Example 11 (2), and monoclonal antibodies are prepared as in Example 11 (3).

[0207] Similarly, DNA encoding the amino acid sequence shown in Fig.13a - Fig.13d or the sequence located downstream of the disintegrin domain of the above sequence is prepared, and inserted into an expression vector pEFBOS, followed by transformation of C2 cells by the resulting vector. The transformant is treated as in Example 11 (1), and used as an antigen for immunization of animals such as mouse, rat and hamster. Antibodies recognizing human Meltrin  $\gamma$  is screened as in Example 11 (2), and monoclonal antibodies are prepared as in Example 11 (3).

SEQUENCE LISTING

#### [0208]

20

25

30

35

40

45

50

55

- (1) GENERAL INFORMATION:
- (i) APPLICANT:
  - (A) NAME: MOCHIDA PHARMACEUTICAL CO., LTD.
  - (B) STREET: 7, Yotsuya 1-chome, Shinjuku-ku
  - (C) CITY: Tokyo
  - (E) COUNTRY: Japan
  - (F) POSTAL CODE (ZIP): 160
  - (ii) TITLE OF INVENTION: MELTRINS
  - (iii) NUMBER OF SEQUENCES: 28
  - (iv) COMPUTER READABLE FORM:
    - (A) MEDIUM TYPE: Floppy disk
    - (B) COMPUTER: IBM PC compatible
    - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
  - (v) CURRENT APPLICATION DATA:

APPLICATION NUMBER: EP 96935358.0

SEQ ID NO:1

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 6915 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: cDNA
  - (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO (vii) IMMEDIATE SOURCE

(B) CLONE: JM109(pBSMel α)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

	GCCA	GAGT.	AG (	CGCGC	CGCGC	es ca	ACGC	ACACI	A CA	CGGG	GAGG	GGA	GAAA	GTT '	ITTT.	<b>PTTGAA</b>	60
10	AAAA!	<b>IGAA</b>	AG (	GCTAC	ACTO	ig Ci	rgct	CAGC	ACC	CCGG	<b>SCGC</b>	TGC	GCGA(	GGG (	GGTC	<b>ECGGCA</b>	120
	GACT	CAGG	GC 2	AGTA(	GAC1	T C	ccca	AGCTY	GGG	CGCC	CGCG	TGG	<b>SATG</b>	CTG (	CAGC	CTGGC	180
	CGCGG	3GGC	CC (	CCGA	GCAG	C TO	<b>SCAC</b> (	CCA(	GC(	CGGC	BACA	ATG	GCA	GAG	CGC	CCG	235
												Met	Ala	Glu	Arg	Pro	
15	GCG (	CGG (	CGC	GCG	CCC	ccc	GCC	CGC	GCC	CTC	CTG	CTG	GCC	CTG	GCT	GGG	283
	Ala	Arg .	Arg	Ala	Pro	Pro	Ala	Arg	Ala	Leu	Leu	Leu	Ala	Leu	Ala	Gly	
	GCC (	erra (	CTG	GCG	CCC	CGT	GCA	GCC	CGA	GGG	ATG	AGT	TTG	TGG	GAC	CAG	331
	Ala 1																
20																	
	AGA (	3GA (	gÇt	TAC	GAA	GTG	GCC	AGA	GCC	TCC	CIT	CTG	AGC	AAG	GAC	CCT	379
	Arg (	3ly 2	Ala	Tyr	Glu	Val	Ala	Arg	Ala	Ser	Leu	Leu	Ser	Lys	Asp	Pro	
	GGG 1	ATC (	CCA	GGA	CAG	AGC	ATC	CCA	GCC	AAG	GAT	CAT	CCA	GAC	GTG	CTG	427
<i>2</i> 5	Glv 1	Ile :	Pro	Glv	Gln	Ser	Ile	Pro	Ala	Lvs	Asp	His	Pro	GEA	Val	Leu	

	ACT	GTG	CAA	CTG	CAG	CTG	GAG	AGC	CGA	GAC	CTG	ATC	CTC	AGC	CTG	GAA	475
	Thr	Val	Gln	Leu	Gln	Leu	Glu	Ser	Arg	Asp	Leu	Ile	Leu	Ser	Leu	Glu	
5			_													CTG	523
	Arg	Asn	Glu	Gly	Leu	Ile	Ala	Asn	Gly	Phe	Thr	Glu	Thr	His	Tyr	Leu	
						~~~		~~~				~~~		~~m	~~ m	mam.	
						GTC											571
10	Gin	дая	GIY	Thr	Авр	Val	261	nea	TILL	Arg	ABII	uis	Int	мвр	пта	Cya	
10	ጥአሮ	TAC	СУТ	CCA	ראיד	GTG	CAA	CCA	CAT	CCT	GCA	TCA	GTG	GTC	AGC	CTC	619
						Val											<b></b>
	-,-	-3-		,													
	AGT	ACT	TGC	TCT	GAT	CTC	CGG	GGA	CTT	ATC	ATG	TTT	GAA	AAT	AAA	ACG	667
15	Ser	Thr	Сув	Ser	Asp	Leu	Arg	Gly	Leu	Ile	Met	Phe	Glu	Asn	Lys	Thr	
						ATG											715
	Tyr	Ser	Leu	Glu	Pro	Met	Lys	Asn	Thr	Thr	Asp	Ser	Tyr	Lys	Leu	Val	
20	CCN	CCT	CNG	NGC.	እጥና	ACG	አልሮ	атс	CAA	ccc	CTYC	T/2T	GGG	ተርል	CAG	САТ	763
20	-					Thr										_	,03
		7,4	014							,		-,-	,				
	AAC	AAG	TCC	AAC	CTC	ACC	ATG	GAA	GAT	GTC	TCC	CCT	GGA	ACC	TCT	CAA	811
	Asn	Lys	Ser	Asn	Leu	Thr	Met	Glu	Asp	Val	Ser	Pro	Gly	Thr	Ser	Gln	
25																	
																TAC	859
	Met	Arg	Ala	Arg	Arg	His	Lys	Arg	GIU	Thr	Leu	rys	Met	Thr	rys	Tyr	
	ርም እ	GNG	CTG	CTT	A TVT	GTG	CCA	GAC	AAC	AGA	GAG	بلملمك	CAG	AGG	CAA	GGA	907
30																Gly	
																•	
																CAC	955
	Lys	Asp	Leu	Glu	Lys	Val	Lys	Gln	Arg	Leu	Ile	Glu	Ile	Ala	Asn	His	
														~~~			
35						AGA											1003
	Val	АВР	гув	Pne	lyr	Arg	PIO	Leu	ABII	116	Arg	116	Val	rea	Val	GIY	
	GTG	GAA	GTG	TGG	AAT	GAC	ATC	GAC	AAA	TGC	TCT	ATA	AGC	CAG	GAC	CCA	1051
						Asp											
40				_													
						GAG											1099
	Phe	Thr	Arg	Leu	His	Glu	Phe	Leu	qaA	Trp	Arg	Lys	Ile	Lys	Leu	Leu	
	~~~	003		mac	CNC	GAC	N N TP	CCT	CRG	CTT	NTC.	ncT	ccc	CTT	тат	ישיי	1147
45						Asp											1117
45		9												_	•		
	CAA	GGA	ACC	ACC	ATC	GGC	ATG	GCA	CCC	ATC	ATG	AGC	ATG	TGC	ACT	GCA	1195
	Gln	Gly	Thr	Thr	Ile	Gly	Met	Ala	Pro	Ile	Met	Ser	Met	Сув	Thr	Ala	
							_										
50						GTT											1243
	Glu	Gln	Ser	GTA	GIÀ	Val	val	met	ASP	H15	ser	qaa	ser	Pro	ren	GTÅ	
	GCC	CCD	GTG	אכיכי	ጥባሃጋ	GCA	ሮልጥ	GAG	Calc	GGC	CAC	AAC	TTC	GGG	ATG	AAC	1291
						Ala											
55		, w								,				•			
	CAT	GAC	ACA	CTG	GAG	AGG	GGC	TGC	AGC	TGC	AGA	ATG	GCC	GCA	GAG	AAA	1339

	His	Asp	Thr	Leu	Glu	Arg	Gly	Сув	Ser	Суз	Arg	Met	Ala	Ala	Glu	Lys	
	GGA	GGC	TGC	ATC	ATG	AAC	CCG	TCC	ACG	GGG	TTC	CCA	TTC	CCC	ATG	GTG	1387
5																Val	
																GGC	1435
	Phe	Ser	Ser	Cys	Ser	Arg	Lys	Asp	Leu	Glu	Ala	Ser	Leu	Glu	Lys	Gly	
	ATG	ദേദ	PATA:	TCC	CTC	TTC	AAC	CTA	CCA	GAG	GTC	AAG	CAG	GCC	TTT	GGG	1483
10																Gly	
		•		_													
					GGA												1531
15	GIÀ	Arg	Lys	Cys	GTA	ASI	GIA	ıyr	Val	GIU	GIU	GIÀ	GIU	GIU	Сув	Asp	
15	TGC	GGA	GAA	CCG	GAG	GAA	TGC	ACG	AAT	CGC	TGC	TGT	AAC	GCT	ACC	ACC	1579
	Сув	Gly	Glu	Pro	Glu	Glu	Cys	Thr	Asn	Arg	Сув	Сув	Asn	Ala	Thr	Thr	
					<b>~~</b>	~~~		~~~	<b>7</b>	~~~	CAC	ccc	CBC	<b>TCC</b>	TV:T	GAA	1627
20																Glu	102/
20	-			-		-											
					AAG												1675
	Asp	Сув	Gln	Leu	Lys	Pro	Pro	Gly	Thr	Ala	Сув	Arg	Gly	Ser	Ser	Asn	
25	TCC	TGT	GAC	CTC	CCA	GAA	TTC	TGC	ACA	GGG	ACT	GCC	CCT	CAC	TGT	CCA	1723
					Pro												
																0.000	
																GGT Gly	1771
30	ATA	ABII	val	ıyı	Den	пів	wpp	GLY	N40	710	Cyb	G.I.	GLY	Val	nop	ULY	
																ACG	1819
	Tyr	Cys	Tyr	neA	Gly	Ile	Суз	Gln	Thr	His	Glu	Gln	Gln	Cys	Val	Thr	
	ርፐር	тес	GGA	CCA	GGT	GCT	AAA	CCG	GCT	CCT	GGC	ATC	TGC	TTT	GAG	CGA	1867
35																Arg	
					GGA Gly												1915
	val	ASI	Ser	MIG	GIY	ASP	PIO	ıyı	GIY	NO!!	cys	GLY	Dyo	æp	561	2,3	
40					AAA												1963
	Ser	Ala	Phe	Ala	Lys	Cys	Glu	Leu	Arg	Asp	Ala	Lys	Сув	Gly	Lys	Ile	
	CAG	ጥርም	CAA	CGT	GGT	GCA	AGC	CGA	ССТ	GTC	ATT	GGT	ACC	AAT	GCT	GTT	2011
	Gln	Cys	Gln	Gly	Gly	Ala	Ser	Arg	Pro	Val	Ile	Gly	Thr	Asn	Ala	Val	
45		_															
					AAT Asn												2059
	ser	me	GIU	THE	ASH	116	PIO	GIII	GIII	GIU	GIY	GLY	AL 9	116	Deu	Cyb	
					GTG												2107
50	Arg	Gly	Thr	His	Val	Tyr	Leu	Gly	Asp	Asp	Met	Pro	Asp	Pro	Gly	Leu	
	GTY	Calmin	GCA	CCD	ACA	DAG	ብረታጥ -	GCA	GAD	GGA	444	ATC	TGC	CTC	AAT	CGT	2155
					Thr												
				-													
55					ATC												2203
	Arg	Сув	Gln	Asn	Ile	Ser	Val	Phe	Gly	val	His	rys	Сув	Ala	met	GIN	

	TGC	CAC	GGC	CGA	GGG	GTA	TGT	AAC	AAC	AGG	AAG	AAT	TGC	CAC	TGT	GAA	2251
	Сув	His	Gly	Arg	Gly	Val	Сув	Asn	Asn	Arg	Lys	Asn	Сув	His	Сув	Glu	
5	_															AGC	2299
	Ala	His	Trp	Ala	Pro	Pro	Phe	Сув	Asp	Lys	Phe	Gly	Phe	Gly	Gly	Ser	
	202	CNC	200	CCT	ccc	አጥሮ	NCC.	CAA	CCN	CRT.	220	CNC	ccc	T-13-7	) CT	C-TP N	2347
						ATC Ile											234/
10	1111	vob	261	GLY	710	116	AL 9	GAIL	W.	vob	no	91	GLY	Deu	****	val	
	GGA	ATC	CTG	GTG	AGC	ATC	CTG	TGT	CTG	CTT	GCT	GCT	GGA	TTT	GTG	GTG	2395
						Ile											
								_									
45						ACG											2443
15	Tyr	Leu	Lys	Arg	Lys	Thr	Leu	Met	Arg	Leu	Leu	Phe	Thr	His	Lys	Lys	
		3.00	<b>3</b> ma	<i>-</i>		~~~	300	mom	~m~			maa	~~~			100	2401
						CTA Leu											2491
	1111	****	ne c	GIU	Lys	200	~~3	Cys	Val	,,,,,		561	AL 9			JCI	
20	GGC	CCT	CAC	CTT	GGC	CAG	GCT	CAC	CAC	ACC	ccc	GGG	AAA	GGC	CTG	CTG	2539
	Gly	Pro	His	Leu	Gly	Gln	Ala	His	His	Thr	Pro	Gly	Lys	Gly	Leu	Leu	
						CAT											2587
25	Met	Asn	Arg	ALA	PTO	His	Pne	ASN	Thr	Pro	гÀв	Авр	Arg	HIB	ser	rea	•
	222	TGC	CAG	AAC	ATG	GAC	ATC	AGC	AGG	CCC	CTC	GAC	GCT	CGA	GCC	GTC	2635
						Asp											
	•	•				-			_			-		_			
	CCA	CAG	CTT	CAG	TCA	CCT	CAG	CGA	GTG	CTC	CTG	CCT	CTC	CAC	CAG	ACC	2683
30	Pro	Gln	Leu	Gln	Ser	Pro	Gln	Arg	Val	Leu	Leu	Pro	Leu	His	Gln	Thr	
	-			000		000	~~~			~~~		000	~~~			CC2	2221
						GGC											2731
	210	~ <u>-</u>	<i>_</i>		501	01,			~~3		204		74.4			A24	
35	GTC	AGG	CAG	GCC	CAG	GGC	ATT	CGA	AAA	CCC	AGT	CCT	CCT	CAG	AAG	CCT	2779
	Val	Arg	Gln	Ala	Gln	Gly	Ile	Arg	Lys	Pro	Ser	Pro	Pro	Gln	Lys	Pro	
							_										
						CTG											2827
40	Leu	Pro	ALA	Asp	PLO	Leu	ser	Arg	Thr	ser	Arg	Leu	Thr	ser	Ala	Leu	
	GTG	AGG	ACC	CCA	GGG	CAG	CAG	GAA	CCT	GGG	CAC	CGC	CCA	GCC	ccc	ATC	2875
						Gln											
		-			-					-		_					
45																TAT	2923
45	Arg	Pro	Ala	Pro	Lys	His	Gln	Val	Pro	Arg	Pro	Ser	His	Asn	Ala	Tyr	
	NTC.	220	ጥሮአር	23 3 (2)	ירא מ	CCCA	GBCC	ים מיו	CCTC	מחמי	CTC	מבאת תב	CAG	מאמי	ALAIAS	יאר	2979
	Ile		1 GAC	ww.uc	.CA G		WACC	.6 61	CCIC		GIG	IAMO A	LAG	AAG I		-AC	2313
		-,,,															
50	TATO	TTC	AGC 1	CCAT	TGGA	G TI	GTTG	TTGI	. ACC	AACT	TTC	CGAG	TTTC	TA A	AGTO	TTTAA	3039
	AACA	CCAT	TC 1	CTCC	AGAC	C C1	GGAG	CCAC	TGC	CATO	:GGT	GCTG	TGCI	GT G	GTGC	TTTGT	3099
																CAGGG	3159
																rcgtgc	3219
55																rgtita rggtac	3279 3339
	CITI	CTAT	TTC A	MGGC	CTTA.	rt. C.C.	AAAL	MIAG	CIC	.ccca	CCI	TCCC	.AAGC	r I	IIA	COINC	3337

```
CAGACACACA GCTCAGGACA CCCCAGGGAG AACCTGGCAT GGGTTTTCTT TGTTTGCTTT
     CATTITATCT TITATATTIT GGTATCCCTA TCTTGGGTTG TAGCCAGGGC CTTCAGGAAG 3459
     GTCTTGGGCC ACTGCATGCT AATGGCCTTC AGGTCCTGCA CCCTGAAGCT CTCAGACAAC
5
     AAGTAGGATC TGCTTTCTAG CCAGCAGCTT TGGAGAGAAC CTGGGGTACT GAAAAGAAGG 3579
     TTTGGGGTGT GGTTATACCA GGATGGAGAC TGGAATCCTA ATCTGGGCAA ACATCTGACC 3639
     TTGAGCTGAG CAGCCATGAG CACCTCTAGG AAGCAAGGAC GGCTGAGGTG CTGCACAAGG 3699
     CTCTGCTTTG AGAGCTGGCA GGGGCTTCTC TCTGGCTGCC CTTTGCAGAG TGCTAGCTGG 3759
     CATGGCATGT TGTTTACATC GGGAACAGTG GTGTTTCTAC AAGAAAGCCA CTGCCTGGGC 3819
10
     ACTGCAGACC TCCGTCTCCT GCCCATTTAG AGCTAAGCAA ATTACCACAT TGTCTTCTGG 3879
     ACTGTAATAC AATGACCCTG TGTTCTGACA GATAGAGGAG GCTTTCTATG GAACCATAAC 3939
     TATTTCANA TGTGAACTAG TAACCAGATC TAGTCGATCA ACTCTGGAGA TAGAAATCTC 3999
     CTTTTTACTG CAAGGCTCGA CTTATTAAAA ATTAGGCAGA ATCCATATGC TTGCAAAAGC 4059
     TATAACCACG TGGAATGCTC TTCTCATGGC ACAGCCTGAG TCTGGTATCC TTATTAGTAG 4119
15
     CCATTGGACA AAGCACCCAA AGTTACCTGT GTGTTCTCTT CAAGGCATCC TAATTTCTTC 4179
     AGCATAGAGA GACTCGGTCT TCCTCACATT CTGAACATAC CTATCAATGA CTAAGNCAGC 4239
     AAGGCAATCC GTTTCCGAAT ACTGAGTTGC TCACGGNAAG GCAACCTCAG CCCAGGNAAA 4299
     CTTTTTTCCT CTGNTCTTTC AGTATGTGAC TGGGGAGCTA CCTTCAGAAG CAAATTTTCA 4359
     AGGTGGNCTC AACCCCATNG GATGAAAGNT ATTTTTTTAA AAAATAATTA ATGGTAATGC 4419
     CAGAGGGCTT TCCTGGCNTC CAGATNGGGG CGTAGGNTTG ACTAGCTTTC ACGACAGAAG 4479
20
     GTAAATGACA GCAGTCCTCT ACCTCGTCTG ACTGCTTTAA GATCAAGGCT TCTTTGGAAG 4539
     GGTAACTAAC ATTAATGGCT GGCCTGTGCC TTGAAGCAGA AGGGAAAATA CAGATAAGGA 4599
     ATTTGGTTTG CTTTCTAGAA TCCAAAACTG TATCCAGCAT TGGGAAGCAT GGTCTTCATG 4659
     ACTGGGTAAA TAAATCCACG TCACAGATGC ATAAAAGAAT AACTCTTATG ACATGCCTCT 4719
     TTTTGTGGCA CAGAGACAAT ATTGCTGCCA CTGAGATGCA TACAAAATTT CTGTAACTGA 4779
25
     TATGTCATTC AGTAGTTGTA TTAAGGCCAA ACATCCACAA CTGTAAAGAC TTATAGAGTT
                                                                      4839
     GTGTGGGCGT TGTCTTGTGA GACACACAAA GCCTCAGCTG AAGCGTATGA GCTCCTCCTC
                                                                     4959
     GGGGGGTGCA GAGAGAGTGT GGTTTAGAGG AAGTTGGAGC CATGATCTTC TGCCATCTCC
     CCAGTGTCCA CTAAGGATGC CGATGGTGCC TTACCAGCTG TGCAGTGCTG GCTGCTTGCT
                                                                      5079
     TTTACAGAGC CATGCATTCA TTTCTGAATA AGAACATATT TAATCCTGAA ATTCCCTTAC
     AGGACAGACA GTGTTACTAA AGGAATTCCT CTAAGATACA GTAGTTGTCA ATTAAAGCAT
     ATTTAGCAGT AACTTCAATT TTAACAAAAT TGGGACCCAA TAGCCAGCAT GAGGGTTCTT
     TGACAGAGGG TAGTITCTCT CTCCCTTTCT CCATCCTTCA AATGACAAGA CGTCAAAACT
     AATACAGTTC ATTTGCAGTC CATCTCATGC TTATACATAC TAGAGGTATG ACTAAAGTTG
     GTTGAGTCAT GGGAGACCAT CCCTGAGAAA GTCCAGTCGG TCAAGAGCCT TGCCAGGTGG
     CGTGGCTGGA CGTCCTCCTT TTGTTCCTGC ACTGAGGAAT AGTTATAGGT TATGTGACCC
     CACTTCACAG GCAAGTGGGA GGCGAACCTT GCAGGCATGC CCCTTAAAAG CTGGTCTCAG
     ACCTACAATA GTCCTGAGTC TGTTTTCCCA GCACACAGAG AGCAACAATG CAGTTTTCCA
     TTTCAAAATA TGCATGCCGA GTTTGCGCTC TGTGTGAGTG TTTCCAGGTT ACACATATGG
     GATGACATCA CAGAAACCAC ACAAGCAACA AATTAAATTC TACGGGAAGA AATCCTCCTG
     ACTGGTCTCT GAGGAGACAT TTTTATGCCT TCTTAACTTT ATTAGGAACT CTCAGGCTGA
     AGCTAGGGGT CATTGTCCCC CAACAAATCA ATACAAAGCC ATCAATGNAC TCTCGAAGAA
     CTGCCAAACC CTGATCTGTG TGAATGTTCT CAGGAGCCTG TGATCCCCAT GGTGCTANAA
     AGAGGCTGGA GCTGGGCCAA CAAGAAGGCC TAAGAGTCCT CCTGCCTCTC AGCAGATGTT
45
     TACTGAGCAC TCTGAGCCAG AAGCACCCCG ACAACCAGGA GGACGATNGC TGGGCAGTAG
     GGCGCCCAGC CACTTGCAGC TCTTTCCTCT GAGGCCCGCT TTGTGTTTTA ATTCCCTTCT
     GTCAGGCCCC AANCAGNGGA CACTGTCCTA TAGACCTCCC TCTNAGTTTT CAGACGGCCT
                                                                      6159
     AAGCCATACA CAAATGCCCC AGACTAAGAA ACACCAATAC NTCCCAGCAG TCCCCAAGAA
     CTGGTTTTTA AACACTATGA CAAGTAGAAG AGGGTGTCAC AGAGGCCATT TTTTTTCTTT
50
     TCTTTCCACT CATACTGGAA CCTAGGTCCT CTCTCTACAC TCCTAGTTCC TTTACACAAC 6339
     TCGGCAGTGG CTCCATTACA CCAAGGACAC AGAAAAACAC AGGTACCGAT TTGCCTTCCT 6399
     CTCCTGCCAA TCACAAGTGC CTTACTCTGA CCAGACCCAT GACAAAACCT CTGTCATCCA 6459
     AGAGAGCCAA CTCTCTACCT TTGTTACTAC TTCAAGCCAA TGTGGTAACT GCTAACCTTC 6519
     AAGGGTCACC TAAACAGTAT AGTCCAACCT TCACCAGGAC CATAGCACAG AGCAACCTCC 6579
55
     AGNACACAC CACACACAC CCTTGAATCT ATCCCACAGC ATATCAACCC ACAGTGACCT
                                                                       6639
```

6759
6819
6879
6915

# 10 SEQ ID NO:2:

5

15

20

25

30

35

40

45

50

55

# (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 903 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

31 ·

Met Ala Glu Arg Pro Ala Arg Arg Ala Pro Pro Ala Arg Ala Leu Leu Leu Ala Leu Ala Gly Ala Leu Leu Ala Pro Arg Ala Ala Arg Gly Met Ser Leu Trp Asp Gln Arg Gly Ala Tyr Glu Val Ala Arg Ala Ser Leu Leu Ser Lys Asp Pro Gly Ile Pro Gly Gln Ser Ile Pro Ala Lys Asp His Pro Asp Val Leu Thr Val Gln Leu Gln Leu Glu Ser Arg Asp Leu Ile Leu Ser Leu Glu Arg Asn Glu Gly Leu Ile Ala Asn Gly Phe Thr Glu Thr His Tyr Leu Gln Asp Gly Thr Asp Val Ser Leu Thr Arg Asn His Thr Asp His Cys Tyr Tyr His Gly His Val Gln Gly Asp Ala Ala Ser Val Val Ser Leu Ser Thr Cys Ser Asp Leu Arg Gly Leu Ile Met Phe Glu Asn Lys Thr Tyr Ser Leu Glu Pro Met Lys Asn Thr Thr Asp Ser Tyr Lys Leu Val Pro Ala Glu Ser Met Thr Asn Ile Gln Gly Leu Cys Gly Ser Gln His Asn Lys Ser Asn Leu Thr Met Glu Asp Val Ser Pro Gly Thr Ser Gln Met Arg Ala Arg Arg His Lys Arg Glu Thr Leu Lys Met Thr Lys Tyr Val Glu Leu Val Ile Val Ala Asp Asn Arg Glu Phe Gln Arg Gln Gly Lys Asp Leu Glu Lys Val Lys Gln Arg Leu Ile Glu Ile Ala Asn His Val Asp Lys Phe Tyr Arg Pro Leu Asn Ile Arg Ile Val Leu Val Gly Val Glu Val Trp Asn Asp Ile Asp Lys Cys Ser Ile Ser Gln Asp Pro Phe Thr Arg Leu His Glu Phe Leu Asp Trp Arg Lys Ile Lys Leu Leu Pro Arg Lys Ser His Asp Asn Ala Gln Leu Ile

40

5

10

15

20

**2**5

*3*0

*3*5

45

50

5

10

15

20

25

30

*3*5

40

45

50

55

Ser Gly Val Tyr Phe Gln Gly Thr Thr Ile Gly Met Ala Pro Ile Met Ser Met Cys Thr Ala Glu Gln Ser Gly Gly Val Val Met Asp His Ser Asp Ser Pro Leu Gly Ala Ala Val Thr Leu Ala His Glu Leu Gly His Asn Phe Gly Met Asn His Asp Thr Leu Glu Arg Gly Cys Ser Cys Arg Met Ala Ala Glu Lys Gly Gly Cys Ile Met Asn Pro Ser Thr Gly Phe Pro Phe Pro Met Val Phe Ser Ser Cys Ser Arg Lys Asp Leu Glu Ala Ser Leu Glu Lys Gly Met Gly Met Cys Leu Phe Asn Leu Pro Glu Val Lys Gln Ala Phe Gly Gly Arg Lys Cys Gly Asn Gly Tyr Val Glu Glu Gly Glu Glu Cys Asp Cys Gly Glu Pro Glu Glu Cys Thr Asn Arg Cys CVB Asn Ala Thr Thr Cys Thr Leu Lys Pro Asp Ala Val Cys Ala His Gly Gln Cys Cys Glu Asp Cys Gln Leu Lys Pro Pro Gly Thr Ala Cys Arg Gly Ser Ser Asn Ser Cys Asp Leu Pro Glu Phe Cys Thr Gly Thr Ala Pro His Cys Pro Ala Asn Val Tyr Leu His Asp Gly His Pro Cys Gln Gly Val Asp Gly Tyr Cys Tyr Asn Gly Ile Cys Gln Thr His Glu Gln Gln Cys Val Thr Leu Trp Gly Pro Gly Ala Lys Pro Ala Pro Gly Ile Cys Phe Glu Arg Val Asn Ser Ala Gly Asp Pro Tyr Gly Asn Cys Gly Lys Asp Ser Lys Ser Ala Phe Ala Lys Cys Glu Leu Arg Asp Ala Lys Cys Gly Lys Ile Gln Cys Gln Gly Gly Ala Ser Arg Pro Val Ile Gly Thr Asn Ala Val Ser Ile Glu Thr Asn Ile Pro Gln Glu Gly Gly Arg Ile Leu Cys Arg Gly Thr His Val Tyr Leu Gly Asp Asp Met Pro Asp Pro Gly Leu Val Leu Ala Gly Thr Lys Cys Ala Glu Gly Lys Ile Cys Leu Asn Arg Arg Cys Gln Asn Ile Ser Val Phe Gly Val His Lys Cys Ala Met Gln Cys His Gly Arg Gly Val Cys Asn Asn Arg Lys Asn Cys His Cys Glu Ala His Trp Ala Pro Pro Phe Cys Asp Lys Phe Gly Phe Gly Gly Ser Thr Asp Ser Gly Pro Ile Arg Gln Ala Asp Asn Gln Gly Leu Thr Val Gly Ile Leu Val Ser Ile Leu Cys Leu Leu Ala Ala Gly Phe Val Val Tyr Leu Lys Arg Lys Thr Leu Met Arg Leu Leu Phe Thr His Lys Lys Thr Thr Met Glu Lys Leu Arg Cys Val His Pro

Ser Arg Thr Pro Ser Gly Pro His Leu Gly Gln Ala His His Thr Pro Gly Lys Gly Leu Leu Met Asn Arg Ala Pro His Phe Asn Thr Pro Lys Asp Arg His Ser Leu Lys Cys Gln Asn Met Asp Ile Ser Arg Pro Leu Asp Ala Arq Ala Val Pro Gln Leu Gln Ser Pro Gln Arg Val Leu Leu 10 Pro Leu His Gln Thr Pro Arg Ala Pro Ser Gly Pro Ala Arg Pro Leu Pro Ala Ser Pro Ala Val Arg Gln Ala Gln Gly Ile Arg Lys Pro Ser pro Pro Gln Lys Pro Leu Pro Ala Asp Pro Leu Ser Arg Thr Ser Arg 15 Leu Thr Ser Ala Leu Val Arg Thr Pro Gly Gln Gln Glu Pro Gly His Arg Pro Ala Pro Ile Arg Pro Ala Pro Lys His Gln Val Pro Arg Pro 20 Ser His Asn Ala Tyr Ile Lys 698

#### SEQ ID NO:3:

25

30

35

45

50

55

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 6345 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: not relevant
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vii) IMMEDIATE SOURCE
  - (B)CLONE: JM109(pBSMel β)
- 40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

	GGC	cece	GGC	AGGC	DTAA	GC A	reece:	ATGT	G TG	ATTG	CGGA	CAG'	rgagi	AGG (	SCCG1	<b>TTGCTA</b>	60
5	TC						GGC (										107
							CCG Pro										155
10							GGT Gly										203
15							TCA Ser										251
							GTC Val										299
20							CAC His										347
25	٠																
<i>30</i>																	
35																	
40																	
<b>4</b> 5														•			
50																	

	TGC	TAC	ACT	GCA	agt	GGC	AAT	CCT	CAA	ACC	AGC	ACG	CTG	AAG	TCT	GAG	395
	Сув	Tyr	Thr	Ala	Ser	Gly	Asn	Pro	Gln	Thr	Ser	Thr	Leu	Lys	Ser	Glu	
5	GAT	CAC	TGC	TTT	TAC	CAC	GGG	ACT	GTG	AGG	GAC	GTG	GAT	GAG	TCC	AGT	443
	Asp	His	Cys	Phe	Tyr	His	Gly	Thr	Val	Arg	Авр	Val	Asp	Glu	Ser	Ser	
																AGA	491
10	Val	Thr	Leu	Ser	Thr	Сув	Arg	Gly	Ile	Arg	Gly	Leu	Ile	Ile	Val	Arg	
	AGT	AAC	CTC	AGC	TAC	ATC	ATC	GAG	CCC	GTC	CCT	AAC	AGC	GAC	AGC	CAA	539
	Ser	Asn	Leu	Ser	Tyr	Ile	Ile	Glu	Pro	Val	Pro	Asn	Ser	qaA	Ser	Gln	
	CAC	CGT	ATT	TAC	AGA	TCC	GAA	CAT	CTC	ACG	CTG	CCC	CCG	GGG	AAC	TGT	587
15														_		Сув	
																TTT	635
	Gly	Phe	Glu	His	Ser	Gly	Pro	Thr	Ser	Lys	qaA	Trp	Ala	Leu	Gln	Phe	
20	ACA	CAT	CAG	ACC	AAA	AAG	CAA	CCT	CGC	AGA	ATG	AAA	CGG	GAA	GAT	CTA	683
	Thr	His	Gln	Thr	Lys	Lys	Gln	Pro	Arg	Arg	Met	Lys	Arg	Glu	Asp	Leu	
	CAC	TCT	ATG	AAG	TAC	GTG	GAG	CTT	TAC	CTG	GTG	GCT	GAT	TAT	GCA	GAG	731
25	His	Ser	Met	Lys	Tyr	Val	Glu	Leu	Tyr	Leu	Val	Ala	Авр	Tyr	Ala	Glu	
	TTT	CAG	AAG	AAT	CGA	CAT	GAC	CAG	GAT	GCC	ACC	AAA	CGC	AAG	CTC	ATG	779
	Phe	Gln	Lys	Asn	Arg	His	Авр	Gln	qaA	Ala	Thr	Lys	Arg	Lys	Leu	Met	
	GAG	ATT	GCC	AAC	TAT	GTT	GAT	AAG	TTT	TAC	CGC	TCC	CTG	AAC	ATC	CGA	827
30																Arg	
	ATT	GCA	CTT	GTC	GGC	TTG	GAG	GTG	TGG	ACG	CAT	GGG	GAT	AAG	TGT	GAA	875
	Ile	Ala	Leu	Val	Gly	Leu	Glu	Val	Trp	Thr	His	Gly	qaA	Lys	Cys	Glu	
<i>3</i> 5	Carr	TCA	GAG	λът	CCC	TAC	тст	ACC	CTC	TCC	TCC	ملسلسل	CTT	AGT	TGG	AGG	923
						Tyr											,,,,
	CGC	AAG	CTG	CTT	GCT	CAG	AAG	AGC	CAT	GAC	AAT	GCT	CAG	CTA	ATC	ACG	971
40	Arg	Lys	Leu	Leu	Ala	Gln	Lys	Ser	His	Asp	Asn	Ala	Gln	Leu	Ile	Thr	
	GGC	AGG	TCC	TTC	CAA	GGC	ACC	ACC	ATT	GGC	CTG	GCC	CCC	CTC	ATG	GCC	1019
	Gly	Arg	Ser	Phe	Gln	Gly	Thr	Thr	Ile	Gly	Leu	Ala	Pro	Leu	Met	Ala	
	ATG	TGC	TCC	GTG	TAC	CAG	TCT	GGA	GGA	GTT	AGC	ATG	GAC	CAC	TCC	GAG	1067
45						Gln											
	AAT	GCC	ATT	GGT	GTA	GCC	TCC	ACT	GTG	GCC	CAT	GAG	ATT	GGC	CAC	AAC	1115
						Ala											
50	بلحليل	GGC	ATY:	<b>DCC</b>	СЪТ	GAT	тст	GCA	CAC	TGC	тст	TCT	GCC	AGT	GCA	GCC	1163
						Asp											
		•				-										AAA	1211
						Met											1211
55	vah	GIÀ	GIA	cyn	116	MEL	wia	~	MA		G <sub>1</sub> y	*****			-10	-73	

	GTG	TTC	AGT	TGG	TGT	AAC	AGG	AAG	GAG	CTG	GAC	AGG	TAT	CTG	CAG	ACA	1259
	Val	Phe	Ser	Trp	Сув	neA	Arg	Lys	Glu	Leu	Asp	Arg	Tyr	Leu	Gln	Thr	
5	GGA	GGA	GGG	ATG	TGT	CTC	TCC	AAC	ATG	CCG	GAC	ACT	AGG	ACG	CTG	TAT	1307
-	Gly	Gly	Gly	Met	Сув	Leu	Ser	Asn	Met	Pro	Asp	Thr	Arg	Thr	Leu	Tyr	
	GGA	GGC	CGG	AGG	TGT	GGC	AAC	GGG	TAC	CTG	GAA	GAC	GGT	GAA	GAA	TGT	1355
	Gly	Gly	Arg	Arg	Сув	Gly	Asn	Gly	Tyr	Leu	Glu	Asp	Gly	Glu	Glu	Сув	
10																	
						GAG Glu											1403
	wab	Cys	GIY	GIU	GIU	GIU	GIU	Cyb	270	ADII	710	cys	cys	ABII	ALG	GEL	
																TGC	1451
15 .	Asn	Сув	Thr	Leu	Lys	Glu	Gly	Ala	Glu	Сув	Ala	His	Gly	Ser	Сув	Сув	
	CAC	CAG	TGC	AAG	CTG	GTG	GCT	CCT	GGA	ACC	CAG	TGT	CGG	GAG	CAG	GTT	1499
	His	Gln	Сув	Lys	Leu	Val	Ala	Pro	Gly	Thr	Gln	Сув	Arg	Glu	Gln	Val	
20	CGG	CAA	TGT	GAC	CTC	CCC	GAG	TTC	TGC	ACC	GGC	AAG	TCT	ccc	CAC	TGC	1547
						Pro											
	000	200	220	ጥአጥ	ሞአጥ	CNG	እጥር	CNT	ccc	ACC	ccc	TGC	GNG	CCT	GGC	CAG	1595
																Gln	1333
25																	
						GGC											1643
	Ala	Tyr	Cys	Tyr	Asn	GIA	Met	Сув	Leu	TOF	171	GIN	GIU	GIN	Сув	Gln	
	CAG	CTG	TGG	GGA	CCT	GGA	GCC	CGG	CCT	GCC	CTC	GAT	CTT	TGC	TTT	GAG	1691
30	Gln	Leu	Trp	Gly	Pro	Gly	Ala	Arg	Pro	Ala	Leu	<b>Asp</b>	Leu	Сув	Phe	Glu	
	»GG	GTG	ממ	GCT	CCT	GGT	GAC	ACC	ТАТ	a so	AAC	тст	GGC	DAG	GGC	TTG	1739
						Gly											
	_					_											
<i>3</i> 5						AAG Lys											1787
	ASII	Gry	GIII	-7-	~9	пyо	Cys	561		g	mp	744	LyC	CID	2020	2,5	
						ACC											1835
	Ile	Gln	Сув	Gln	Ser	Thr	Gln	Ala	Arg	Pro	Leu	Glu	Ser	Asn	Ala	Val	
40	TCT	ATT	GAC	ACC	ACC	ATC	ACC	TTG	AAC	GGG	AGG	CGG	ATC	CAC	TGT	CGG	1883
	Ser	Ile	qaA	Thr	Thr	Ile	Thr	Leu	Asn	Gly	Arg	Arg	Ile	His	Cys	Arg	
	GGC	ACC	CAC	GTC	TAC	CGG	GGT	CCT	GAG	GAG	GAG	GAA	GGG	GAA	GGT	GAC	1931
45						Arg											
															~~~		
						CTG Leu											1979
		u	- L		,					,		-,-	-,-			<del></del>	
50						GGG											2027
	His	Ile	Сув	Phe	Glu	Gly	Gln	Cys	Arg	Asn	Thr	Ser	Phe	Phe	GIU	Thr	
	GAA	GGC	TGT	GGG	AAA	AAG	TGC	AAT	GGC	CAC	GGG	GTC	TGC	AAC	AAC	AAC	2075
						Lys											
55	224	224	est of the second	ሮአጥ	TCC	TTC	رست	CCC	TOG	TOT.	CCA	ር ር	ጥጥር	ጥርጥ	ממ	ACC	2123
	~~~	MAC	191	~~1	100	***		55	100	101		1			~~~		~ . 4 3

	Lys	Asn	Сув	His	Cys	Phe	Pro	Gly	Trp	Ser	Pro	Pro	Phe	Сув	Asn	Thr	
	CCG	GGA	GAT	GGT	GGC	AGC	GTC	GAC	AGT	GGT	CCT	TTG	CCC	CCT	AAG	AGT	2171
5						Ser											
	GTG	GGT	CCC	GTG	ATC	GCT	GGG	GTG	TTT	TCA	GCT	CTC	TTC	GTG	TTG	GCA	2219
	Val	Gly	Pro	Val	Ile	Ala	Gly	Val	Phe	Ser	Ala	Leu	Phe	Val	Leu	Ala	
10	GTT	CTG	GTG	CTA	CTG	TGT	CAC	TGC	TAC	AGA	CAG	AGC	CAC	AAA	CTG	GGC	2267
	Val	Leu	Val	Leu	Leu	Суз	His	Сув	Tyr	Arg	Gln	Ser	His	Lys	Leu	Gly	
	AAA	CCC	TCG	GCT	CTC	CCT	TTC	AAG	CTG	CGG	CAT	CAG	TTC	AGT	TGT	CCC	2315
	Lys	Pro	Ser	Ala	Leu	Pro	Phe	Lys	Leu	Arg	His	Gln	Phe	Ser	Сув	Pro	
15		300	CT1	T-C-T	CNG	AGT	com	CCA	n cm	000	C N TP	ecc	220	CCA	N.CT	There's	2363
						Ser											2363
	1110						,	,		1							
						CAG										_	2411
20	Lys	Leu	Gln	Thr	Pro	Gln	Gly	Lys	Arg	Lys	Val	Thr	Asn	Thr	Pro	Glu	
	TCC	CTC	CGG	AAG	CCG	TCC	CAC	CCC	CCT	CTC	CGG	ccc	CCT	CCA	GAC	TAC	2459
	Ser	Leu	Arg	Lys	Pro	Ser	His	Pro	Pro	Leu	Arg	Pro	Pro	Pro	qaA	Tyr	
	~~~	000	CTOTO	CNN	TO C	CCA	COT	CCN	CCN	under.	TV-C	CCN	C እ ጥ	CALC	220	NCG	2507
25						Pro											2307
		9														<b>3</b>	
						CCA											2555
	Ala	Ala	Gly	Ser	Ser	Pro	Glu	Ala	Gly	Ala	Arg	Ile	Glu	Arg	Lys	Glu	
30	TCA	GCC	AGG	AGG	ССТ	CCC	CCA	AGC	CGA	CCC	ATG	ccc	CCT	GCA	CCT	AAC	2603
						Pro											
						GAC Asp											2651
35	cys	Deu	Deu	Jei	<b>0111</b>				5		3			· · · ·	-7-		
						GTG											2699
	Leu	Pro	Ala	Asn	Pro	Val	Pro	Gly	Gln	Arg	Thr	Gly	Pro	Arg	Ser	Gly	
	GGC	ACC	TCC	СТС	СТТ	CAG	CCC	CCT	ACT	TCT	GGT	CCT	CAG	ccc	CCC	AGG	2747
40						Gln											
																	2795
						GTT Val											2/35
	110	FIQ	AL G	VA.2	7.0			_,,				-,-					
45						AGC				TAG	LAGTO	TC G	AGAA	GTT	C		2842
	Val	Gly	Ala	Ile	Ile	Ser	Ser	Lys	Ile								
	TTGT	TCC	AT C	GAAC	ACTO	C GG	ATGO	CATO	GAA	GGTC	CAG	AAGA	AAGA	CG C	CTT	TCACC	2902
	CATO	CTG	LAG (	TTTC	GCAG	CT	TCTO	GAAC	GTC	CCTC	CATC	CCC	GAAT	CT C	CCTI	CTTAC	2962
50																TAAGTG	3022
																ngggaa ngagct	3082 3142
																LGGAAC	3202
																TACTG	3262
55																CATGT	3322

	AATAAGCCAT	GCTCCCCTCC	CCTGCCTTTC	TTCACATTCC	CACTCCCATA	TTTACACGGG	3382
						GGTGGGGTGG	3442
					• • • • • • • • • • • • • • • • • • • •	TAGAATCCCC	3502
5						TTCCATGTGT	3562
5			<del>-</del>			TATGGCAGTC	3622
	•					ATATGGGGTC	3682
	•••••					TGTATGTTTG	3742
						TGTGTGTGTG	3802
10						CTGCATGTAC	3862
10							
						CTCATATCTA	3922
						TTGGCATCCT	3982
						TCTGAGGGGG	4042
15						TGGGCAGGGT	4102
15				-		CCTGGTTGCT	4162
						GGTGGACCTT	4222
	<del>-</del> · ·		ATTTGGAACA				4282 4342
						TCTCTAGGGG	4402
20						GAAGAGGCAA	4462
20	•					AGTAAGCTTT	4522
			GGTGACCGTG			ATCTCAGGAA	4582
						CAGGCCACCC	4582
						TATCAGAAAA	4702
25	•					TACTGTGCTA	4762
23			TATTCTGTGA				4822
	•	- <del>-</del>	ACACACAACT				4882
			TCCCTTCCCA				4942
						TGAAAGAATA	5002
30						ACACCAAGTC	5062
-		•		*		TGAGCAAGAA	5122
						GAAGATGTTA	5182
						CGTTTAGTGG	5242
						GCCTGGTCTA	5302
35				-		CTCTAGAAAA	5362
			TCATGTATCG				5422
			TACAGACTGC				5482
	TGGGTTCAAA	CACAGCCCTT	CATATGTATA	TAGCCAAACA	CTCTACAACT	GAGCTACATC	5542
	CTCCAGCCTA	GGCTGTAAAT	GTTTTTTGGA	GCTAGATTAG	CTGCCTGCCA	ACCTTAGAAC	5602
40	TGCAAAGCCA	TTCCTGACCT	GTAAACCTCA	GCTCTCCATC	TCTATAAGAG	GTATAGCCTG	5662
	GGCTAATACC	GTCCAAGTTA	CAACTCCTTG	CTTGCTTTCT	<b>GTTCCTTCTA</b>	GCCTTGGTGA	5722
	CTTCCACCAG	GAAGAGAATA	CCCCCTCTCT	ACCCCTGCTC	CAAGACACTG	TAGATGCTAG	5782
	TGTCGGAGTG	TTCTCTGTAA	CGCGACAGTT	CCTTCTGTTG	CAATAGCCCC	CCTGCAACAC	5842
	TGCAATAATC	CTTCAGTGTC	TCCCCTGGGC	TCAATTCACT	TCCTTATTTG	ACAAAGTGGA	5902
45	GGTGAGACTT	GTATTCTTAA	AATTGGAGGC	TAGTTATTTT	GTCAAATGCA	TGTAATGAAC	5962
	AGACCCGAAG	GAATCCTCCA	CACACAAGCC	AGGGAACACC	AACTGGAAAG	GTACCCCGTC	6022
	CCAGGGAAGC	CTGCTAGGGA	GAGGTTCTGT	AGAATCCGAG	CCTAGCACCC	CAAAGTCATG	6082
						GCAAATGTGA	6142
						ATGTCCTCCT	6202
50						ATGGTGCCTT	6262
	GTTTTTTGTT	GTTGTTGTTA	TTTTTTTCTC	CTTGTTTGTA	AAATTAATT	ACAAATTGTC	6322
	ATGAGGAAAA	алалалала	AAA				6345

55 SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 920 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein

5

10

15

20

25

30

35

40

45

50

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Pro Gly Arg Ala Gly Val Ala Arg Phe Cys Leu Leu Ala Leu Ala Leu Gln Leu His Trp Pro Leu Ala Ala Cys Glu Pro Gly Trp Thr Thr Arg Gly Ser Gln Glu Gly Ser Pro Pro Leu Gln His Glu Leu Ile Ile Pro Gln Trp Arg Thr Ser Glu Ser Pro Gly Arg Gly Lys His Pro Leu Arg Ala Glu Leu Arg Val Met Ala Glu Gly Arg Glu Leu Ile Leu Asp Leu Glu Lys Asn Glu His Leu Phe Ala Pro Ala Tyr Thr Glu Thr Cys Tyr Thr Ala Ser Gly Asn Pro Gln Thr Ser Thr Leu Lys Ser Glu Asp His Cys Phe Tyr His Gly Thr Val Arg Asp Val Asp Glu Ser Ser Val Thr Leu Ser Thr Cys Arg Gly Ile Arg Gly Leu Ile Ile Val Arg Ser Asn Leu Ser Tyr Ile Ile Glu Pro Val Pro Asn Ser Asp Ser Gln His Arg Ile Tyr Arg Ser Glu His Leu Thr Leu Pro Pro Gly Asn Cys Gly Phe Glu His Ser Gly Pro Thr Ser Lys Asp Trp Ala Leu Gln Phe Thr His Gln Thr Lys Lys Gln Pro Arg Arg Met Lys Arg Glu Asp Leu His Ser Met Lys Tyr Val Glu Leu Tyr Leu Val Ala Asp Tyr Ala Glu Phe Gln Lys Asn Arg His Asp Gln Asp Ala Thr Lys Arg Lys Leu Met Glu Ile Ala Asn Tyr Val Asp Lys Phe Tyr Arg Ser Leu Asn Ile Arg Ile Ala Leu Val Gly Leu Glu Val Trp Thr His Gly Asp Lys Cys Glu Val Ser Glu Asn Pro Tyr Ser Thr Leu Trp Ser Phe Leu Ser Trp Arg Arg Lys Leu Leu Ala Gln Lys Ser His Asp Asn Ala Gln Leu Ile Thr Gly Arg Ser Phe Gln Gly Thr Thr Ile Gly Leu Ala Pro Leu Met Ala Met Cys Ser Val Tyr Gln Ser Gly Gly Val Ser Met Asp His Ser Glu Asn Ala Ile Gly Val Ala Ser Thr Val Ala His Glu Ile Gly His Asn Phe Gly Met Ser His Asp Ser Ala His Cys Cys Ser Ala Ser Ala Ala Asp Gly Gly Cys Ile Met Ala Ala Ala Thr Gly His Pro Phe Pro Lys Val

5

5

Phe	Ser	Trp	Cys	Asn	Arg	Lys	Glu	Leu	Asp	Arg	Tyr	Leu	Gln	Thr	Gly
Gly	Gly	Met	Сув	Leu	Ser	Asn	Met	Pro	qaA	Thr	Arg	Thr	Leu	Tyr	Gly
Gly	Arg	Arg	Cys	Gly	Asn	Gly	Tyr	Leu	Glu	Asp	Gly	Glu	Glu	Сув	Asp
Сув	Gly	Glu	Glu	Glu	Glu	Cys	Lys	Asn	Pro	Сув	Cys	ùsA	Ala	Ser	Asr
Сув	Thr	Leu	Lys	Glu	Gly	Ala	Glu	Сув	Ala	His	Gly	Ser	Сув	Cys	His
Gln	Cys	Lys	Leu	Val	Ala	Pro	Gly	Thr	Gln	Сув	Arg	Glu	Gln	Val	Arg
Gln	Сув	Asp	Leu	Pro	Glu	Phe	Сув	Thr	Gly	Lys	Ser	Pro	His	Сув	Pro
Thr	Asn	Tyr	Tyr	Gln	Met	Asp	Gly	Thr	Pro	Сув	Glu	Gly	Gly	Gln	Ala
Tyr	Сув	Tyr	Asn	Gly	Met	Сув	Leu	Thr	Tyr	Gln	Glu	Gln	Cys	Gln	Glr
Leu	Trp	Gly	Pro	Gly	Ala	Arg	Pro	Ala	Leu	Asp	Leu	Сув	Phe	Glu	Arg
Val	Asn	Ala	Ala	Gly	Asp	Thr	Tyr	Gly	Asn	Сув	Gly	Lys	Gly	Leu	Asr
Gly	Gln	Tyr	Arg	Lys	Сув	ser	Pro	Arg	Asp	Ala	Lys	Сув	Xaa	Lys	Ile
Gln	Сув	Gln	Ser	Thr	Gln	Ala	Arg	Pro	Leu	Glu	Ser	Asn	Ala	Val	Ser
Ile	Asp	Thr	Thr	Ile	Thr	Leu	Asn	Gly	Arg	Arg	Ile	His	Сув	Arg	Gly
Thr	His	Val	Tyr	Arg	Gly	Pro	Glu	Glu	Glu	Glu	Gly	Glu	Gly	Asp	Met
Leu	Asp	Pro	Gly	Leu	Val	Met	Thr	Gly	Thr	Lys	Сув	Gly	His	Asn	Hie
Ile	Cys	Phe	Glu	Gly	Gln	Cys	Arg	Asn	Thr	Ser	Phe	Phe	Glu	Thr	Glu
Gly	Cys	Gly	Lys	Lys	Сув	Asn	Gly	His	Gly	Val	Сув	neA	Asn	Asn	Lys
Asn	Сув	His	Cys	Phe	Pro	Gly	Trp	Ser	Pro	Pro	Phe	Cys	Asn	Thr	Pro
Gly	qeA	Gly	Gly	Ser	Val	Asp	Ser	Gly	Pro	Leu	Pro	Pro	Lys	Ser	Val
Gly	Pro	Val	Ile	Ala	Gly	Val	Phe	Ser	Ala	Leu	Phe	Val	Leu	Ala	Val
Leu	Val	Leu	Leu	Сув	His	Сув	Tyr	Arg	Gln	Şer	His	Lys	Leu	Gly	Lys
Pro	Ser	Ala	Leu	Pro	Phe	Lys	Leu	Arg	His	Gln	Phe	Ser	Сув	Pro	Phe
Arg	Val	Ser	Gln	Ser	Gly	Gly	Thr	Gly	His	Ala	neA	Pro	Thr	Phe	Lys
Leu	Gln	Thr	Pro	Gln	Gly	Lys	Arg	Lys	Val	Thr	Aen	Thr	Pro	Glu	Sex
Leu	Arg	Lys	Pro	Ser	His	Pro	Pro	Leu	Arg	Pro	Pro	Pro	Asp	Tyr	Lev
Arg	Val	Glu	Ser	Pro	Pro	Ala	Pro	Leu	Ser	Ala	His	Leu	Asn	Arg	Ala
Ala	Gly	Ser	Ser	Pro	Glu	Ala	Gly	Ala	Arg	Ile	Glu	Arg	Lys	Glu	Ser

	A	la 3	Arg	Arg	Pro	Pro	Pro	Ser	Arg	Pro	Met	Pro	Pro	Ala	Pro	Asn	Сув
5	L	eu I	Leu	Ser	Gln	Asp	Phe	Ser	Arg	Pro	Arg	Pro	Pro	Gln	Lys	Ala	Leu
J	P	co A	\la	Asn	Pro	Val	Pro	Gly	Gln	Arg	Thr	Gly	Pro	Arg	Ser	Gly	Gly
	T	ar S	Ser	Leu	Leu	Gln	Pro	Pro	Thr	Ser	Gly	Pro	Gln	Pro	Pro	Arg	Pro
10	P	co A	<b>ila</b>	Val	Pro	Val	Pro	Lys	Leu	Pro	Glu	Tyr	Arg	Ser	Gln	Arg	Val
	G:	Ly A	Ala	Ile	Ile	Ser	Ser	Lys	Ile 716								
15																	
	SEQ ID NO	):5:															

20

25

30

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 3928 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: not relevant
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vii) IMMEDIATE SOURCE
- (B)CLONE: JM109(pBSMel γ)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

35

40

45

50

	GTT	GCAA	GGA :	<b>I</b> GAC	CGAA	GC G(	GAGG	CGGC	G GC	CGCG	CGTI	GAG	CGGA	ACC :	TGCC	GAAGC	C 60
5	CTC					CGC (											108
																GGG Gly	156
10																ATT Ile	204
																AGT Ser	252
15																ATT Ile	300
20						AAC Asn										GTT Val	346
																GTA Val	396
25																TCC Ser	444
	GCG	GTT	GCT	GTG	AGC	GCC	TGC	TTT	GGA	CTC	AGA	GGC	TTG	CTG	CAT	TTG	492
30																	

5

	Ala	Val	Ala	Val	Ser	Ala	Суз	Phe	Gly	Leu	Arg	Gly	Leu	Leu	His	Leu	
	GAG	AAT	GCC	AGT	TTT	GGA	ATT	GAA	CCT	CTG	CAC	AAC	AGC	TCA	CAC	TTT	540
5																Phe	310
-	~~~				<b></b>												
																AGA Arg	588
	0.0			•	-,-			p	01,			42	<b>414</b>		Deu	my.	
10																GAT	636
	Сув	Gly	Val	Ser	Asn	Arg	qaA	Thr	Glu	Lys	Glu	Gly	Thr	Gln	Gly	Asp	
	GAG	GAG	GAG	CAT	CCG	AGT	GTC	ACT	CAG	CTG	CTG	CGC	AGA	AGA	AGA	GCT	684
						Ser											
15				~~~													
																AAG Lys	732
	<b>V</b> 42	200					-,-						***	Vu.	, asp	-Jy -S	
	-															GAG	780
20	Glu	Arg	Tyr	Asp	Met	Met	Gly	Arg	asn	Gln	Thr	Ala	Val	Arg	Glu	Glu	
	ATG	ATT	CGC	TTA	GCA	AAC	TAC	CTG	GAT	AGC	ATG	TAC	ATC	ATG	TTA	AAC	828
	Met	Ile	Arg	Leu	Ala	Asn	Tyr	Leu	Asp	Ser	Met	Tyr	Ile	Met	Leu	Asn	
				cm2	~~~			cen.	<i>a</i>				~~~				454
25						GTT Val											876
							1							3			
						GGA											924
20	Ile	Asn	Ile	Ile	GIA	Gly	Ala	GIÀ	qaA	Val	Leu	Gly	Asn	Phe	Val	Gln	
30	TGG	CGG	GAA	AAG	TTC	CTT	ATA	ACT	CGT	CGG	AGA	CAC	GAC	AGT	GCA	CAG	972
	Trp	Arg	Glu	Lys	Phe	Leu	Ile	Thr	Arg	Arg	Arg	His	Asp	Ser	Ala	Gln	
	TOTAL	C TTT	ሞማሂደ	מממ	מממ	GGC	Jelek	CCT	GGA	ልርጥ	CCA	CGA	NTC	ccc	ملحليك	CTA	1020
35						Gly											1020
•				•	-	_		_	•			•					
						AGG											1068
	GIY	Thr	vaı	Сув	ser	Arg	ser	HIS	ALA	GIA	GIY	TIE	ASN	vaı	Pne	GIA	
40	CAA	ATC	ACT	GTG	GAG	ACA	TTT	GCA	TCC	ATT	GTT	GCT	CAT	GAA	TTG	GGG	1116
	Gln	Ile	Thr	Val	Glu	Thr	Phe	Ala	Ser	Ile	Val	Ala	His	Glu	Leu	Gly	
	СУТ	מממ	CTT	CCD	STG	AAT	CAT	CAT	GAT	GGG	AGA	GAG	ጉርም	TTC	ጥርጥ	CCA	1164
						Asn											2203
45																	
						ATG Met											1212
	,,,,,	Lys	JCI	Cyu				-	O.,	~~~	Jer	OLY	<b>361</b>	~~y	A011	FIIC	
						GAG											1260
50	Ser	Ser	Сув	Ser	Ala	Glu	Asp	Phe	Glu	Lys	Leu	Thr	Leu	Asn	Lув	Gly	
	GGA	AGC	TGC	CTG	CTT	AAC	ATC	CCG	AAG	ССТ	GAC	GAA	GCC	TAC	AGC	GCG	1308
						Asn											
55						AAG Lys											1356
		عناب	-70	1	~	7					1		314	-70	ų.	-10	

	GGC	ACA	GCG	AAG	GAG	TGT	GAG	GTG	GAC	CCA	TGC	TGT	GAA	GGA	AGC	ACT	1404
	Gly	Thr	Ala	Lys	Glu	Сув	Glu	Val	qaA	Pro	Сув	Сув	Glu	Gly	Ser	Thr	
5																AAA	1452
	-	-							Сув		-		_	_			
									TCC								1500
10	_	-					_	-	Ser		-	_					
									AAC Asn								1548
	GIU	Сув	Asp	VAI	PIO	GIU	ıyı	Сув	MBII	GIA	261	Ser	GIII	FIIC	Сув	PIO	
	CCA	GAT	GTC	TTC	ATT	CAG	AAT	GGA	TAT	CCT	TGC	CAG	AAC	AGC	AAA	GCC	1596
15	Pro	Asp	Val	Phe	Ile	Gln	Asn	Gly	Tyr	Pro	Сув	Gln	Asn	Ser	Lys	Ala	
																~~~	
									TAT							Val	1644
	Tyr	Сув	Tyr	ASII	GIY	nec	Сув	GIII	ıyı	Tyt	Asp	ALG	GIII	Cyb		VGI	
20	ATC	TIT	GGT	TCA	AAG	GCT	AAG	GCT	GCC	CCA	AGA	GAT	TGC	TTC	ATT	GAA	1692
	Ile	Phe	Gly	Ser	Lys	Ala	Lys	Ala	Ala	Pro	Arg	Asp	Сув	Phe	Ile	Glu	
									GGC								1740
<i>2</i> 5	vaı	ABn	ser	гув	GIY	wab	wrg	PHE	Gly	Yell	Cys	GIA	FILE	261	GIY	361	
	GAG	TAC	AAG	AAG	TGT	GCC	ACT	GGG	AAC	GCG	CTG	TGT	GGA	AAG	CTT	CAA	1788
	Glu	Tyr	Lys	Lys	Сув	Ala	Thr	Gly	Asn	Ala	Leu	Сув	Gly	Lys	Leu	Gln	
																200	
30									GTG Val								1836
	Сув	GIU	ABII	val	GIII	мар	rie L	PIO	441	PME	GIY	116	VAI	710	<i>~~</i>	116	
	ATT	CAG	ACA	CCC	AGT	CGA	GGC	ACC	AAA	TGC	TGG	GGT	GTG	GAT	TTC	CAG	1884
	Ile	Gln	Thr	Pro	Ser	Arg	Gly	Thr	Lys	СЛа	Trp	Gly	Val	qaA	Phe	Gln	
<i>3</i> 5	- Carrier	com	maa	CNC	CONT	CCN	CNC	CCA	GGG	בערת	CTY2	አአጥ	GAA	cac	ACC	222	1932
									Gly								2,3,2
		1							•					•		•	
									AAT								1980
40	Сув	Asp	Ala	Gly	Lys	Ile	Сув	Arg	Asn	Phe	Gln	Сув	Val	Asn	Ala	Ser	
	GTC	CTG	ТАА	TAT	GAC	TGT	GAC	ATT	CAG	GGA	AAA	TGT	CAT	GGC	CAT	GGG	2028
																Gly	
						-											
45									CAC								2076
	Val	Сув	neA	ser	Asn	Lys	Asn	Cys	His	Сув	GIU	ABD	GIÀ	Trp	AIA	PIO	
	CCA	CAC	TGT	GAC	ACC	AAA	GGA	TAT	GGA	GGA	AGC	GTG	GAC	AGC	GGG	CCG	2124
									Gly								
50			_	-		_	-										
50									CTG								2172
	Thr	Tyr	Asn	Ala	Lув	Ser	Thr	ALA	Leu	arg	qea	GTÅ	Leu	ren	vai	rne	
	TTC	TTC	CTA	ATC	GTC	CCC	CTT	GTT	GCG	GCT	GCC	ATT	TTC	CTC	TTT	ATC	2220
									Ala								
55																	

				GAA Glu													2268
5				AGA Arg													2316
	AGT	ATC	TCC	AGA Arg	CCA	CCA	GGG	GGC	CCA	AAT	GTC	TCC	AGA	CCA	CCA	GGG	2364
10	GGC	CCA	GGT	GTC	TCC	AGA	CCA	CCA	GGG	GGC	CCA	GGT	GTC	TCC	AGA	CCA	2412
	CCA	GGG	GGC	Val CCA	GGT	GTC	TCC	AGA	CCG	CCA	CCT	GGG	CAT	GGA	AAC	AGA	2460
15	TTC	CCA	GTA	Pro CCA	ACC	TAC	GCC	GCC	AAG	CAG	CCT	GCG	CAG	TTC	CCG	TCA	2508
20				Pro CCA													2556
20	Arg	Pro	Pro	Pro	Pro	Gln	Pro	Lys	Ile	Ser	Ser	Gln	Gly	Asn	Leu	Ile	2601
25	Pro	Ala	Arg	Pro	Ala	Pro	Ala	Pro	Pro	Leu	Tyr	Ser	Ser	Leu	Thr		2661
	GTTC	STTT?	TT :	TTTT ACAAJ	CTGI	AT GT	PTTT(	TAAC	A AAJ C ACJ	AGCC1	TTC CCT	TCTT AGTC	rccaj eggj	ACC A	ATGA SCGAJ	ATGITT ATGAAC AACACA	2721 2781
30	CTGC	CCAT	TTT (	STGGA CAAA?	ATTTA OAAT1	LA TO	CACI IATI	TGAC	GTG	GATI	DAAT DTDT	TTAT ACT	TCT(	DAE	LATG!	PTTTCT PTACTG STAATC STATTA	2841 2901 2961 3021
	ACTI	raggi Atato	AAG J	ACTA!	ATTGC IGTCI	C AJ	)AATA	GTC1	C GCZ	ATTTT BAAGO	CAT CAGA	TATO	CATO	SGA ?	TTAA( BAGC	CAGCCA PTACAC FACTCA	3081 3141 3201
35	GTT	ATTG(	CT :	rccat Ptate	GGC1	TA TO	SATC!	TTCA TCA	A ACT	TATAI TCT1	CAA	TTAT AACT	KTADI LATDI	AGA I	AATC( ACTA:	EATTTA IGACAG IGTTGT	3261 3321 3381
40	TTTC	TTTC	EAA (	STGCA SGAGO	CACT	C TA	ATGG? AATA(	TACGI	A GGT	GTT?	TAGT TGAG	ATAC GTGT	CCA.	AGC J	AGATA	AGGTGT CTTGAG CTCTAG	3441 3501 3561
	CATT	TTTI	CA T	ITTAC SAGAJ	TAAC AAGGA	G TO	STGC?	rggg: Lagt(	CCT	rgtc1 rtcc1	CTT CTAG	TTG/	CTA CTAC	ATA 1	CTTT(	CGTAAA ATACTG ATGATA	3621 3681 3741
<b>4</b> 5	TTG	ACACT	AA I	AATTI IAACI	laati Laadi	C AT	LATT! LTTA:	ATTI AGTTI	TT	Cati Laaai	AATC ATAA	TATT	LATA!	aag j ngt (	AAGT:	TTAATA ATATAC AAAAAA	3801 3861 3921
		AAA															3928

(2) INFORMATION FOR SEQ ID NO:6:

50

55

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 845 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

# (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

5

Į	Met	Gly	Pro	Arg	Ala	Leu	Ser	Pro	Leu	Ala	Ser	Leu	Arg	Leu	Arg	Trp
1	Leu	Leu	Ala	Сув	Gly	Leu	Leu	Gly	Pro	Val	Leu	Glu	Ala	Gly	Arg	Pro
2	Asp	Leu	Glu	Gln	Thr	Val	His	Leu	Ser	Ser	Tyr	Glu	Ile	Ile	Thr	Pro
٠	Irp	Arg	Leu	Thr	Arg	Glu	Arg	Arg	Glu	Ala	Leu	Gly	Pro	Ser	Ser	Gln
(	Gln	Ile	Ser	Tyr	Val	Ile	Gln	Ala	Gln	Gly	Lys	Gln	His	Ile	Ile	His
1	Leu	Glu	Arg	Asn	Thr	qeA	Leu	Leu	Pro	Asn	Авр	Phe	Val	Val	Tyr	Thr
•	ryr	Asp	Lys	Glu	Gly	Ser	Leu	Leu	Ser	Asp	His	Pro	Asn	Val	Gln	Ser
1	His	Cys	His	Tyr	Arg	Gly	Tyr	Val	Glu	Gly	Val	Gl'n	Asn	Ser	Ala	Val
2	Ala	Val	Ser	Ala	Сув	Phe	Gly	Leu	Arg	Gly	Leu	Leu	His	Leu	Glu	Asn
2	Ala	Ser	Phe	Gly	Ile	Glu	Pro	Leu	His	Asn	Ser	Ser	His	Phe	Glu	His
:	Ile	Phe	Tyr	Pro	Met	Asp	Gly	Ile	His	Gln	Glu	Pro	Leu	Arg	Cys	Gly
1	<b>Val</b>	Ser	Asn	Arg	qeA	Thr	Glu	Lys	Glu	Gly	Thr	Gln	Gly	qaA	Glu	Glu
(	31u	His	Pro	Ser	Val	Thr	Gln	Leu	Leu	Arg	Arg		Arg -205		Val	Leu
1	Pro	Gln	Thr	Arg	Tyr	Val	Glu	Leu	Phe	Ile	Val	Val	qeA	Lys	Glu	Arg
•	Гут	Asp	Met	Met	Gly	Arg	Asn	Gln	Thr	Ala	Val	Arg	Glu	Glu	Met	Ile
1	Arg	Leu	Ala	Asn	Tyr	Leu	Asp	Ser	Met	Tyr	Ile	Met	Leu	Asn	Ile	Arg
				Val	_											
			_	Gly												
		-		Leu											•	
				Gly												
		_		Arg												
				Thr												
		•		Asn												
		-		Met												
	_			Glu												
(	Сув	Leu	Leu	Asn	Ile	Pro	Lys	Pro	Авр	Glu	Ala	Tyr	Ser	Ala	Pro	Ser

Сув	Gly	Asn	Lys	Leu	Val	Asp	Pro	Gly	Glu	Glu	Сув	qaA	Сув	Gly	Thr
Ala	Lys	Glu	Сув	Glu	Val	Asp	Pro	Cys	Сув	Glu	Gly	Ser	Thr	Cys	Lya
Leu	Lys	Ser	Phe	Ala	Glu	Cys	Ala	Tyr	Gly	Asp	Сув	Cys	Lys	Двр	Сув
Gln	Phe	Leu	Pro	Gly	Gly	Ser	Met	Cys	Arg	Gly	Lys	Thr	Ser	Glu	Cys
Asp	Val	Pro	Glu	Tyr	Суз	Asn	Gly	Ser	Ser	Gln	Phe	Сув	Pro	Pro	qeA
Val	Phe	Ile	Gln	Asn	Gly	Tyr	Pro	Cys	Gln	Asn	Ser	Lys	Ala	Tyr	Сув
Tyr	Asn	Gly	Met	Сув	Gln	Tyr	Tyr	Asp	Ala	Gln	Сув	Gln	Val	Ile	Phe
Gly	Ser	Lys	Ala	Lys	Ala	Ala	Pro	Arg	Asp	Cys	Phe	Ile	Glu	Val	Asn
Ser	Lys	Gly	Asp	Arg	Phe	Gly	Asn	Cys	Gly	Phe	Ser	Gly	Ser	Glu	Tyr
Lys	Lys	Сув	Ala	Thr	Gly	Asn	Ala	Leu	Cys	Gly	Lys	Leu	Gln	Сув	Glu
Asn	Val	Gln	Asp	Met	Pro	Val	Phe	Gly	Ile	Val	Pro	Ala	Ile	Ile	Gln
Thr	Pro	Ser	Arg	Gly	Thr	Lys	Сув	Trp	Gly	Val	qaA	Phe	Gln	Leu	Gly
Ser	qaA	Val	Pro	Asp	Pro	Gly	Met	Val	Asn	Glu	Gly	Thr	Lys	Сув	Asp
Ala	Gly	Lys	Ile	Сув	Arg	Asn	Phe	Gln	Суз	Val	Asn	Ala	Ser	Val	Leu
Asn	Tyr	Asp	Cys	qaA	Ile	Gln	Gly	Lys	Сув	His	Gly	His	Gly	Val	Сув
Asn	Ser	Asn	Lys	Asn	Сув	His	Cys	Glu	Asp	Gly	Trp	Ala	Pro	Pro	His
Сув	Asp	Thr	Lys	Gly	Tyr	Gly	Gly	Ser	Val	Asp	Ser	Gly	Pro	Thr	Tyr
Asn	Ala	Lys	Ser	Thr	Ala	Leu	Arg	qaA	Gly	Leu	Leu	Val	Phe	Phe	Phe
Leu	Ile	Val	Pro	Leu	Val	Ala	Ala	Ala	Ile	Phe	Leu	Phe	Ile	Lys	Arg
Asp	Glu	Leu	Arg	Lys	Thr	Phe	Arg	Lys	Lys	Arg	Ser	Gln	Met	Ser	Asp
Gly	Arg	Asn	Gln	Ala	Asn	Val	Ser	Arg	Gln	Pro	Gly	Asp	Pro	Ser	Ile
Ser	Arg	Pro	Pro	Gly	Gly	Pro	Asn	Val	Ser	Arg	Pro	Pro	Gly	Gly	Pro
Gly	Val	Ser	Arg	Pro	Pro	Gly	Gly	Pro	Gly	Val	Ser	Arg	Pro	Pro	Gly
Gly	Pro	Gly	Val	Ser	Arg	Pro	Pro	Pro	Gly	His	Gly	Asn	Arg	Phe	Pro
Val	Pro	Thr	Tyr	Ala	Ala	Lys	Gln	Pro	Ala	Gln	Phe	Pro	Ser	Arg	Pro
Pro	Pro	Pro	Gln	Pro	Lys	Ile	Ser	Ser	Gln	Gly	neA	Leu	Ile	Pro	Ala
Arg	Pro	Ala	Pro	Ala	Pro	Pro	Leu	Tyr	Ser	Ser	Leu	Thr 640			

SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

55

5		(E (C	A) LEN B) TYP C) STR D) TOP	E: nuc	leic ac	cid SS: not		ant									
10		(iii) HY	OLECU POTH NTI-SE	HETICA ENSE:	AL: NO NO	)											
15		·	B)CLOI		Ţ			·	10:7:								
20		CCT Pro															48
20		GAG Glu															96
25		CAC His															144
		ATC Ile															192
30	GGT Gly	GCT	AAA Lys	CCT Pro	GCC Ala	CCT Pro	GGG Gly	ATC Ile	TGC Cys	TTT Phe	GAG Glu	AGA Arg	GTC Val	TAA CIBA	TCT Ser	GCA Ala	240
		GAA Glu													_		286
35		TGC Cys															321
40	SEC	O D N	O:8:														
		(i) SE	QUEN	CE CH	HARAC	CTERI	STICS	<b>S</b> :			,						
45		(E	N) LEN B) TYP D) TOP	E: ami	ino aci	id	cids										
			DLECL EQUE		•		N: SE	Q ID N	<b>1</b> 0:8:								
50		, 5															

		Lys 1	Pro	Ala	Gly	Thr 5	Ala	Сув	Arg	Asp	Ser 10	Ser	Asn	Ser	Сув	Asp 15	Leu
5			Glu	Phe	Сув 20		Gly	Ala	Ser	Pro 25		Сув	Pro	Ala	Asn 30		Tyr
		Leu	His	ДВР 35		His	Ser	Сув	Gln 40	Asp	Val	Asp	Gly	Tyr 45	Cys	Xaa	Asn
		Gly	Ile 50		Gln	Thr	His	Glu 55	Gln	Gln	Сув	Val	Thr 60	Leu	Trp	Gly	Pro
10		Gly 65		Lys	Pro	Ala	Pro 70		Ile	Cys	Phe	Glu 75	Arg	Val	Asn	Ser	Ala 80
15		Gly	Glu	. Pro	туг	Gly 85		Cys	Gly	' Lys	Val		Lys	Ser	Ser	Phe 95	
		Lys	Сув	Glu	100		<b>Asp</b>	Ala	Lys	Cys 105		Lys					
20	SEQ ID	NO:9:															
	(i) S	EQUE	NCE	CHAR	ACTE	RISTIC	S:										
25		(A) LE (B) TY (C) ST (D) TC	PE: n	ucleic EDN	acid ESS: n		vant										
30	(iii) (iv)	MOLEC HYPO ANTI-S IMME	THETI SENSE	CAL: I E: NO	NO	,											
35		(B) CL	ONE:	JM10	9 (pBS	huMγ	G238)	)									
	(xi)	SEQU	ENCE	DESC	CRIPTI	ON: S	EQ ID	NO:9	:								
40				-													
45																	
50																	
55																	

																TIT	48
	Ala	Lys	Ser	Сув	Ile	Met	Asn	Ser	Gly	Ala	Ser	Gly	Ser	Arg	Asn	Phe	
5	AGC	AGT	TGC	AGT	GCA	GAG	GAC	TTT	GAG	AAG	TTA	ACT	TTA	AAT	AAA	GGA	96
	Ser	Ser	Сув	Ser	Ala	Glu	qaA	Phe	Glu	Lys	Leu	Thr	Leu	Asn	Lys	Gly	
	GGA	AAC	TGC	CTT	CTT	AAT	ATT	CCA	AAG	CCT	GAT	GAA	GCC	TAT	AGT	GCT	144
10																Ala	
	CCC	TCC	TGT	GGT	AAT	AAG	TTG	GTG	GAC	GCT	GGG	GAA	GAG	TGT	GAC	TGT	192
																Cys	
	GGT	ACT	CCA	AAG	GAA	TGT	GAA	TTG	GAC	ССТ	TGC	TGC	GAA	GGA	AGT	ACC	- 240
15	Gly	Thr	Pro	Lys	Glu	Сув	Glu	Leu	qaA	Pro	САв	Сув	Glu	Gly	Ser	Thr	
																AAA	288
	Сув	Lys	Leu	Lys	Ser	Phe	Ala	Glu	Сув	Ala	Tyr	Gly	Asp	Сув	Сув	Lys	
20																AGT	336
	Asp	Сув	Arg	Phe	Leu	Pro	Gly	Gly	Thr	Leu	Суб	Arg	Gly	Lys	Thr	Ser	
																CAG	384
25	Glu	Cys	yab	Val	Pro	Glu	Tyr	Сув	Asn	Gly	Ser	Ser	Gln	Phe	Cys	Gln	
																GCC	432
	Pro	qeA	Val	Phe	Ile	Gln	Asn	Gly	Tyr	Pro	Сув	Gln	Asn	Asn	Lys	Ala	
																GTC	480
30	Tyr	Сув	Tyr	Asn	Gly	Met	CAa	Gln	Tyr	Tyr	Asp	Ala	Gln	Сув	Gln	Val	
																GAA	528
	Ile	Phe	Gly	Ser	Lys	Ala	Lys	Ala	Ala	Pro	Lys	Asp	Сув	Phe	Ile	Glu	
35	GTG	AAT	TCT	AAA	GGT	GAC	AGA	TTT	GGC	AAT	TGT	GGT	TTC	TCT	GGC	AAT	576
	Val	Asn	Ser	Lvs	Glv	GRA	Arg	Phe	Glv	Asn	Cvs	Glv	Phe	Ser	Gly	Asn	

	GAA	TAC	AAG	aag	TGT	GCC	ACT	GGG	AAT	GCT	TTG	TGT	GGA	AAG	CTT	CAG	624
	Glu	Tyr	Lys	Lys	Cys	Ala	Thr	Gly	Asn	Ala	Leu	Сув	Gly	Lys	Leu	Gln	
5	TGT	GAG	AAT	GTA	CAA	GAG	ATA	CCT	GTA	TTT	GGA	ATT	GTG	CCT	GCT	ATT	672
	САв	Glu	Asn	Val	Gln	Glu	Ile	Pro	Val	Phe	Gly	Ile	Val	Pro	Ala	Ile	
	ATT	CAA	ACG	CCT	AGT	CGA	GGC	ACC	AAA	TGT	TGG	GGT	GTG	GAT	TTC	CAG	720
10	Ile	Gln	Thr	Pro	Ser	Arg	Gly	Thr	Lys	Cys	Trp	Gly	Val	Asp	Phe	Gln	
	CTA	GGA	TCA	GAT	GTT	CCA	GAT	ССТ	GGG	ATG	GTT	AAC	GAA	GGC	ACA	AAA	768
	Leu	Gly	Ser	Asp	Val	Pro	Asp	Pro	Gly	Met	Val	asn	Glu	Gly	Thr	Lys	
15	• • •															TCT	816
.•	Сув	Gly	Ala	Gly	Lys	Ile	Сув	Arg	Asn	Phe	Gln	Сув	Val	Asp	Ala	Ser	
																GGG	864
	Val	Leu	Asn	Tyr	qaA	Сув	qaA	Val	Gln	Lys	Lys	Сув	His	Gly	His	Gly	
20	_					AAG											912
	Val	Сув	Asn	Ser	Asn	Lys	Asn	Сув	His	Сув	Glu	Asn	Gly	Trp	Leu	Pro	
																TCG	960
25	Gln	Ile	Val	Arg	Leu	Lys	qaA	Thr	Arg	Ser	Ser	Leu	ser	Ile	Pro	ser	
	ACC		A				•										967
	Thr	ser			•												

30 (2) INFORMATION FOR SEQ ID NO:10:

*3*5

40

45

50

55

# (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 322 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

	7	lla 1	Lys	Ser	Сув	Ile	Met	Asn	Ser	Gly	Ala 10	Ser	Gly	Ser	Arg	Asn 15	Phe
	s	_	Ser	Суз	Ser	Ala	Glu	Asp	Phe	Glu	Lys	Leu	Thr	Leu	Asn		Gly
5	_	-1			20	• •••		~1.	D	25	D	B.00	<b>61</b>	33.	30	C	31.
	G	зтХ	Asn	35	Leu	Leu	ABR	He	40	гув	Pro	ABD	GIU	45	TYI	ser	AIA
	E	?ro	Ser 50		Gly	Asn	Lys	Leu 55	Val	Asp	Ala	Gly	Glu 60	Glu	Сув	Asp	Cys
10	G	31y 65		Pro	Lys	Glu	Сув 70	Glu	Leu	Asp	Pro	Cys 75	Сув	Glu	Gly	Ser	Thr 80
	C		Lys	Leu	Lys	Ser 85	Phe	Ala	Glu	Сув	Ala 90	Tyr	Gly	Asp	Сув	Сув 95	Lys
15	,	qaA	Сув	Arg	Phe		Pro	Gly	Gly	Thr 105		Сув	Arg	Gly			Ser
15	G	3lu	Сув	_	100 Val	Pro	Glu	Tyr			Gly	Ser	Ser		110 Phe	Cys	Gln
	E	?ro	Asp	115 Val	Phe	Ile	Gln	Asn	120 Gly	Tyr	Pro	Суз	Gln	125 Asn	Asn	Lys	Ala
20																	
			130					135					140				
	1	yr		Tyr	Asn	Gly	Met		Gln	Tyr	Tyr	qeA		Gln	Cys	Gln	Val
ae.	_	45	250	<b>~</b> 1	C	7	150	1	<b>71</b> 0	777	Pro	155 Lve	Non.	~~	Dhe	Tla	160
25				-	Ser	165					170					175	
		•			Lys 180					185					190		
30	G	lu	Tyr	Lys 195	Lys	Сув	Ala	Thr	Gly 200	aaA	Ala	Leu	Сув	Gly 205	Lys	Leu	Gln
	C	:ys		Asn	Val	Gln	Glu		Pro	Val	Phe	Gly		Val	Pro	Ala	Ile
	7	'1e	210 Gln	Thr	Pro	Ser	Ara	215 Glv	Thr	Lvs	Cvs	Trp	220 Glv	Val	Asp	Phe	Gln
	2	25					230	_				235					240
35			_		Asp	245		_			250					255	
	C	ув	Gly	Ala	Gly 260	Lys	Ile	Сув	Arg	Asn 265	Phe	Gln	Сув	Val	Asp 270	Ala	Ser
40	V	al	Leu	Asn 275	Tyr	Asp	Сув	Asp	Val 280	Gln	Lys	Lys	Cys	His 285	Gly	His	Gly
	v		Cys 290		Ser	Asn	Lys	Asn 295		His	Cys	Glu	Asn 300		Trp	Leu	Pro
	G			Val	Arg	Leu	Lys		Thr	Arg	Ser	Ser		Ser	Ile	Pro	Ser
45	_	05	_				310					315					320
	Т	hr	Ser														
	SEQ ID N	0:11	:														
50	(i) SE	QUE	NCE (	CHAR	ACTE	RISTIC	S:										
	(A	A) LE	NGTH	l: 2848	3 base	pairs											
	(E	3) TY	PE: ni	ucleic	acid	•											
55	-	•		DEDNE DGY: II	ESS: n inear	ot rele	vant										

(ii) MOLECULE TYPE: cDNA (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE:	NO
(vii) IMMEDIATE S	SOURCE

5

(B) CLONE:JM109(pMeI  $\alpha$ -25C) JM109 (pMeI  $\alpha$ -26N)

# (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

10	GGG	GAC	CTC	TGG	ATC	CCA	GTG	AAG	AGC	TTC	GAC	TCC	AAG	AAT	CAT	CCA	48
	Gly	Asp	Leu	Trp	Ile	Pro	Val	Lys	Ser	Phe	Asp	Ser	Lys	Asn	His	Pro	
	GAA	GTG	CTG	AAT	ATT	CGA	CTA	CAA	CGG	GAA	AGC	AAA	GAA	CIG	ATC	ATA	96
	Glu	Val	Leu	Asn	Ile	Arg	Leu	Gln	Arg	Glu	Ser	Lys	Glu	Leu	Ile	Ile	
15	224	Citto	CAA	AGA	ДДТ	GDD	CCT	כיזיכ	ATT	acc	AGC	AGT	TTC	ACG	GAA	ACC	144
				Arg													
				CAA	<b>63.6</b>		3.00	CAM	ana	<b>TC</b> C	CTC	CCT	CCA	እልጥ	TAC	»CC	192
~~																	174
20	HIB	TYT	ren	Gln	Asp	GIY	Inr	wab	VAI	ser	Leu	WIG	Arg	7511	-7-	****	
	GGT	CAC	TGT	TAC	TAC	CAT	GGA	CAT	GTA	CGG	GGA	TAT	TCT	GAT	TCA	GCA	240
	Gly	His	Сув	Tyr	Tyr	His	Gly	His	Val	Arg	Gly	Tyr	Ser	Asp	Ser	Ala	

	GTC	AGT	CTC	AGC	ACG	TGT	TCT	GGT	CTC	AGG	GGA	CTT	ATT	GGG	TTT	GAA	288
	Val	Ser	Leu	Ser	Thr	Сув	Ser	Gly	Leu	Arg	Gly	Leu	Ile	Gly	Phe	Glu	
5												~~~					
J																TAC Tyr	336
	Wen	GIU	Ser	171	AGT	Dea	914	710	Met	n y o	Jei	710	1111	ADII	AL 9	-7-	
	AAA	CTC	TTC	CCA	GCG	AAG	AAG	CTG	AAA	AGC	GTC	CGG	GGA	TCA	TGT	GGA	384
10	Lys	Leu	Phe	Pro	Ala	Lys	Lys	Leu	Lys	Ser	Val	Arg	Gly	Ser	Сув	Gly	
10								ama	~~~	CC3	220		~~~		~~	<b>663</b>	433
									_	-						CCA Pro	432
	561										-7-					•••	
15	CCC	TCT	CAG	ACA	TGG	GCA	AGA	AGG	CAT	AAA	AGA	GAG	ACC	CTC	AAG	GCA	480
15	Pro	Ser	Gln	Thr	Trp	Ala	Arg	Arg	His	Lys	Arg	Glu	Thr	Leu	Lys	Ala	
	3 CM	220	TAT	CTC	GNG	CTC	GTYG:	እ <b>ጥ</b> ር	GTG	GCA	GNC	AAC	CGA	GNG	Tala	CAG	528
																Gln	720
200		_	•														
20						CTG											576
	Arg	Gln	Gly	Lys	Asp	Leu	Glu	Lys	Val	Lys	Gln	Arg	Leu	Ile	Glu	Ile	
	сст	ТАА	CAC	GTT	GAC	AAG	TTT	TAC	AGA	CCA	CTG	AAC	ATT	CGG	ATC	GTG	624
																Val	
25																	
																AGT	672
٠	Leu	vaı	GIY	Val	GIU	Val	тър	Asn	Авр	met	мар	гÀв	Сув	Ser	vai	Ser	
	CAG	GAC	CCA	TTC	ACC	AGC	CTC	CAT	GAA	TTT	CTG	GAC	TGG	AGG	AAG	ATG	720
30	Gln	Asp	Pro	Phe	Thr	Ser	Leu	His	Glu	Phe	Leu	Asp	Trp	Arg	Lys	Met	
					~~~		maa.	O . T	C2.0	220	~~~	030	CALINE	CTC	200	CCC	768
																GGG	700
	-1-					-1-											
35																ATG	816
	Val	Tyr	Phe	Gln	Gly	Thr	Thr	Ile	Gly	Met	Ala	Pro	Ile	Met	Ser	Met	
	TGC	ACG	GCA	GAC	CAG	TCT	GGG	GGA	ATT	GTC	ATG	GAC	CAT	TCA	GAC	AAT	864
																Asn	
40																	
						GTG Val											912
	PIO	Leu	GIY	WIG	WIG	VAI	****	Deu	ALG	1110	910	1)CU	GLY	1110	<b>~</b>		
						ACA											960
45	Gly	Met	Asn	His	<b>Asp</b>	Thr	Leu	Asp	Arg	Gly	Сув	Ser	Сув	Gln	Met	Ala	
	CTT	GNG	מממ	CCA	GGC	TGC	ATY	בעדמ	AAC	GCT	ፐርሮ	ACC	ccc	TAC	CCA	للململ	1008
						Сув											2000
			-		_												
50						AGT											1056
	Pro	Met	Val	Phe	Ser	Ser	Сув	5er	Arg	Lys	Asp	Leu	Glu	Thr	Ser	Leu	
	GAG	444	GGA	ATG	GGG	GTG	TGC	CTG	TTT	AAC	CTG	CCG	GAA	GTC	AGG	GAG	1104
						Val											· •
55		-	-		-		-								-		

																GAG	1152
	Ser	Phe	Gly	Gly	Gln	Lys	Сув	Gly	Asn	Arg	Phe	Val	Glu	Glu	Gly	Glu	
5				TGT													1200
	Glu	Сув	Asp	Cys	Gly	Glu	Pro	Glu	Glu	Сув	Met	Asn	Arg	Cys	Сув	Asn	
									~~~	~~		maa		01m		CTC	1040
				TGT													1248
	ALA	Thr	TAF	Сув	Inr	ren	гув	PIO	wsb	ATG	AGT	Cys	ALG	HID	GLY	Deu	
10	TYC	ጥርተ	GAA	GAC	TGC	CAG	CTG	DAG	ССТ	GCA	GGA	ACA	GCG	TGC	AGG	GAC	1296
				Asp													
	0,0	-7-			-,-									•	_	-	
	TCC	AGC	AAC	TCC	TGT	GAC	CTC	CCA	GAG	TTC	TGC	ACA	GGG	GCC	AGC	CCT	1344
15	Ser	Ser	Asn	Ser	Сув	Asp	Leu	Pro	Glu	Phe	Cys	Thr	Gly	Ala	Ser	Pro	
				GCC													1392
	His	Cys	Pro	Ala	Agn	vaı	TYX	Leu	HIB	Asp	GIY	HIS	ser	Сув	GIN	wab	
20	are:	anc	GGC	TAC	TGC	TAC	AAT	GGC	ATC	TGC	CAG	ACT	CAC	GAG	CAG	CAG	1440
20				Tyr													
			1	-1-	-7-	-1-				•							
				CTC													1488
	Сув	Val	Thr	Leu	Trp	Gly	Pro	Gly	Ala	Lys	Pro	Ala	Pro	Gly	Ile	Сув	
25																	
	TTT	GAG	AGA	GTC	AAT	TCT	GCA	GGT	GAT	CCT	TAT	GGC	AAC	TGT	GGC	AAA	1536
	Phe	Glu	Arg	Val	ASD	ser	Ata	GIY	Asp	Pro	Tyr	Grà	ASI	СУВ	GIY	rya	
	CTC	TCC	AAG	AGT	TCC	TTT	GCC	AAA	TGC	GAG	ATG	AGA	GAT	GCT	AAA	TGT	1584
30				Ser													
			-														
																ACC	1632
	Gly	Lys	Ile	Gln	Сув	Gln	Gly	Gly	Ala	Ser	Arg	Pro	Val	Ile	Gly	Thr	
											~~~	~~~	~~~	~~~		666	1680
<i>3</i> 5	AAT	GCC	GTT	TCC	ATA	GAA	The	AAC	TIA	Pro	Len	Gla	Gln	GGA	GIV	CGG Arg	1660
	ASN	AIG	val	Ser	116	GIU	1111	Aou	116	710	Men	GAM	<b>U</b> 111	O <sub>1</sub>	O.J		
	ATT	CTG	TGC	CGG	GGG	ACC	CAC	GTG	TAC	TTG	GGC	GAT	GAC	ATG	CCG	GAC	1728
				Arg													
40																	
	CCA	GGG	CTT	GTG	CTT	GCA	GGC	ACA	AAG	TGT	GCA	GAT	GGA	AAA	ATC	TGC	1776
	Pro	Gly	Leu	Val	Leu	Ala	GIÀ	Thr	rys	Сув	ALA	Asp	GIA	rys	TTE	Cys	•
	caç	እ እ <b>ጥ</b>	ССТ	CAA	TOT	CAA	ДДТ	АТТ	AGT	GTC	TTTT	GGG	GTT	CAC	GAG	TGT	1824
	Leu	AAI	Arg	Gln	CVS	Gln	Asn	Ile	Ser	Val	Phe	Gly	Val	His	Glu	Cys	
45				•	-1-												
				TGC													1872
	Ala	Met	Gln	Сув	His	Gly	Arg	Gly	Val	Сув	Asn	Asn	Arg	Lys	Asn	Сув	
										-	<b>mc</b> -	<b></b>		-	-		1020
50				GCC													1920
	HIB	Сув	GIU	Ala	H18	1.Lb	ATS	PIQ	PEO	rne	cys	wab	ոչ	FNE	GTÅ	PINE	
	GGA	GGA	AGC	ACA	GAC	AGC	GGC	CCC	ATC	CGG	CAA	GCA	GAA	GCA	AGG	CAG	1968
				Thr													
<b>E</b> E	-	-															
55	GAA	GCT	GCA	GAG	TCC	AAC	AGG	GAG	CGC	GGC	CAG	GGC	CAG	GAG	ccc	GTG	2016

# Glu Ala Ala Glu Ser Asn Arg Glu Arg Gly Gln Gly Gln Glu Pro Val

			GCC TCA CTG ACA CTC ATC TGA	2061
5	Gly Ser Gln Glu H	is Ala Ser Thr	Ala Ser Leu Thr Leu Ile *	
	GCCCTCCCAT GACATG	GAGA CCGTGACCA	TGCTGCTGCA GAGGAGGTCA CGCGTCCCCA	2121
	AGGCCTCCTG TGACTG	GCAG CATTGACTC	GTGGCTTTGC CATCGTTTCC ATGACAACAG	2181
	ACACAACACA GTTCTC	GGGG CTCAGGAGG	GAAGTCCAGC CTACCAGGCA CGTCTGCAGA	2241
10	AACAGTGCAA GGAAGG	GCAG CGACTTCCT	GTTGAGCTTC TGCTAAAACA TGGACATGCT	2301
	TCAGTGCTGC TCCTGA	GAGA GTAGCAGGT	ACCACTCTGG CAGGCCCCAG CCCTGCAGCA	2361
	AGGAGGAAGA GGACTC	AAAA GTCTGGCCT	TCACTGAGCC CCCACAGCAG TGGGGGAGAA	2421
	GCAAGGGTTG GGCCCA	GTGT CCCCTTTCC	CAGTGACACC TCAGCCTTGG CAGCCCTGAT	2481
	GACTGGTCTC TGGCTG	CAAC TTAATGCTC	GATATGGCTT TTAGCATTTA TTATATGAAA	2541
15	ATAGCAGGGT TTTAGT	TTTT AATTTATCA	AGACCCTGCC ACCCATTCCA TCTCCATCCA	2601
	AGCAAACTGA ATGGCA	TTGA AACAAACTG	B AGAAGAAGGT AGGAGAAAGG GCGGTGAACT	2661
	CTGGCTCTTT GCTGTG	GACA TGCGTGACCI	GCAGTACTCA GGTTTGAGGG TTTGCAGAAA	2721
	GCCAGGGAAC CCACAG	AGTC ACCAACCCT	CATTTAACAA GTAAGAATGT TAAAAAGTGA	2781
	AAACAATGTA AGAGCC	TAAC TCCATCCCC	GTGGCCATTA CTGCATAAAA TAGAGTGCAT	2841
20	CCCGCCC			2848

#### SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 686 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

30

*3*5

40

45

50

55

25

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Gly Asp Leu Trp Ile Pro Val Lys Ser Phe Asp Ser Lys Asn His Pro Glu Val Leu Asn Ile Arg Leu Gln Arg Glu Ser Lys Glu Leu Ile Ile Asn Leu Glu Arg Asn Glu Gly Leu Ile Ala Ser Ser Phe Thr Glu Thr His Tyr Leu Gln Asp Gly Thr Asp Val Ser Leu Ala Arg Asn Tyr Thr Gly His Cys Tyr Tyr His Gly His Val Arg Gly Tyr Ser Asp Ser Ala Val Ser Leu Ser Thr Cys Ser Gly Leu Arg Gly Leu Ile Gly Phe Glu Asn Glu Ser Tyr Val Leu Glu Pro Met Lys Ser Ala Thr Asn Arg Tyr Lys Leu Phe Pro Ala Lys Lys Leu Lys Ser Val Arg Gly Ser Cys Gly Ser His His Asn Thr Pro Asn Leu Ala Ala Lys Asn Val Phe Pro Pro Pro Ser Gln Thr Trp Ala Arg Arg His Lys Arg Glu Thr Leu Lys Ala Thr Lys Tyr Val Glu Leu Val Ile Val Ala Asp Asn Arg Glu Phe Gln Arg Gln Gly Lys Asp Leu Glu Lys Val Lys Gln Arg Leu Ile Glu Ile Ala Asn His Val Asp Lys Phe Tyr Arg Pro Leu Asn Ile Arg Ile Val

5

5.

10

15

20

25

30

*3*5

40

45

50

55

Leu Val Gly Val Glu Val Trp Asn Asp Met Asp Lys Cys Ser Val Ser Gln Asp Pro Phe Thr Ser Leu His Glu Phe Leu Asp Trp Arg Lys Met Lys Leu Leu Pro Arg Lys Ser His Asp Asn Ala Gln Leu Val Ser Gly Val Tyr Phe Gln Gly Thr Thr Ile Gly Met Ala Pro Ile Met Ser Met Cys Thr Ala Asp Gln Ser Gly Gly Ile Val Met Asp His Ser Asp Asn Pro Leu Gly Ala Ala Val Thr Leu Ala His Glu Leu Gly His Asn Phe Gly Met Asn His Asp Thr Leu Asp Arg Gly Cys Ser Cys Gln Met Ala Val Glu Lys Gly Gly Cys Ile Met Asn Ala Ser Thr Gly Tyr Pro Phe Pro Met Val Phe Ser Ser Cys Ser Arg Lys Asp Leu Glu Thr Ser Leu Glu Lys Gly Met Gly Val Cys Leu Phe Asn Leu Pro Glu Val Arg Glu Ser Phe Gly Gly Gln Lys Cys Gly Asn Arg Phe Val Glu Glu Gly Glu Glu Cys Asp Cys Gly Glu Pro Glu Glu Cys Met Asn Arg Cys Cys Asn Ala Thr Thr Cys Thr Leu Lys Pro Asp Ala Val Cys Ala His Gly Leu Cys Cys Glu Asp Cys Gln Leu Lys Pro Ala Gly Thr Ala Cys Arq Asp Ser Ser Asn Ser Cys Asp Leu Pro Glu Phe Cys Thr Gly Ala Ser Pro His Cys Pro Ala Asn Val Tyr Leu His Asp Gly His Ser Cys Gln Asp Val Asp Gly Tyr Cys Tyr Asn Gly Ile Cys Gln Thr His Glu Gln Gln Cys Val Thr Leu Trp Gly Pro Gly Ala Lys Pro Ala Pro Gly Ile Cys Phe Glu Arg Val Asn Ser Ala Gly Asp Pro Tyr Gly Asn Cys Gly Lys Val Ser Lys Ser Ser Phe Ala Lys Cys Glu Met Arg Asp Ala Lys Cys Gly Lys Ile Gln Cys Gln Gly Gly Ala Ser Arg Pro Val Ile Gly Thr Asn Ala Val Ser Ile Glu Thr Asn Ile Pro Leu Gln Gln Gly Gly Arg Ile Leu Cys Arg Gly Thr His Val Tyr Leu Gly Asp Asp Met Pro Asp Pro Gly Leu Val Leu Ala Gly Thr Lys Cys Ala Asp Gly Lys Ile Cys Leu Asn Arg Gln Cys Gln Asn Ile Ser Val Phe Gly Val His Glu Cys Ala Met Gln Cys His Gly Arg Gly Val Cys Asn Asn Arg Lys Asn Cys His Cys Glu Ala His Trp Ala Pro Pro Phe Cys Asp Lys Phe Gly Phe

GIA	GIA	ser	Thr	Asp	Ser	GIÀ	Pro	Ile	Arg	Gln	Ala	Glu	Ala	Arg	Gln
Glu	Ala	Ala	Glu	Ser	Asn	Arg	Glu	Arg	Gly	Gln	Gly	Gln	Glu	Pro	Val
Gly	Ser	Gln	Glu	His	Ala	Ser	Thr	Ala	Ser	Leu	Thr	Leu	Ile		

_ C	FΩ	חו	NI	7·1	2

	SEQ	ID	NO:	1
10				

5

15

20

/:\	CEOL			CTERIST	10c.
111	うていし	JEINGE	UDADA	ULEDIOL	IUO.

- (A) LENGTH: 394 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

								GAA								TGC Cys	48
25	013	014	0	C, C		C,U	023		424	-	0.14	cyo	7011	<i>-</i>	710	Cys	
	TGC	AAT	GCC	TCT	AAT	TGT	ACC	CTG	AGG	CCG	GGG	GCG	GAG	TGT	GCT	CAC	96
	Сув	Asn	Ala	Ser	Asn	Сув	Thr	Leu	Arg	Pro	Gly	Ala	Glu	Сув	Ala	His	
30	GGC	TCC	TGC	TGC	CAC	CAG	TGT	AAG	CTG	TTG	GCT	CCT	GGG	ACC	CTG	TGC	144
	Gly	Ser	Cys	Cys	His	Gln	Сув	Lys	Leu	Leu	Ala	Pro	Gly	Thr	Leu	Cys	
	CGC	GAG	CAG	GCC	AGG	CAG	TGT	GAC	CTC	CCG	GAG	TTC	TGT	ACG	GGC	AAG	192
	Arg	Glu	Gln	Ala	Arg	Gln	Cys	qeA	Leu	Pro	Glu	Phe	Сув	Thr	Gly	Lys	
35	TCT	CCC	CAC	TGC	CCT	ACC	AAC	TTC	TAC	CAG	ATG	GAT	GGT	ACC	ccc	TGT	240
	Ser	Pro	His	Cys	Pro	Thr	Asn	Phe	Tyr	Gln	Met	qeA	Gly	Thr	Pro	Cys	
	GAG	GGC	GGC	CAG	GCC	TAC	TGC	TAC	AAC	GGC	ATG	TGC	CTC	ACC	TAC	CAG	288
40	Glu	Gly	Gly	Gln	Ala	Tyr	Сув	Tyr	Asn	Gly	Met	Сув	Leu	Thr	Tyr	Gln	
	GAG	CAG	TGC	CAG	CAG	CTG	TGG	GGA	ccc	GGA	GCC	CGA	CCT	GCC	CCT	GAC	336
	Glu	Gln	Сув	Gln	Gln	Leu	Trp	Gly	Pro	Gly	Ala	Arg	Pro	Ala	Pro	qeA	
45	CTC	TGC	TTC	GAG	AAG	GTG	AAT	GTG	GCA	GGA	GAC	ACC	TTT	GGA	AAC	TGT	384
	Leu	Сув	Phe	Glu	Lys	Val	Asn	Val	Ala	Gly	qaA	Thr	Phe	Gly	Asn	Сув	
	GGA	AAG	GAC	A													394
	Gly	Lys	Asp														

SEQ ID NO:14:

50

55

#### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 131 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:
- 5 Gly Glu Glu Cys Asp Cys Gly Glu Glu Glu Glu Cys Asn Asn Pro Cys 1 Cys Asn Ala Ser Asn Cys Thr Leu Arg Pro Gly Ala Glu Cys Ala His Gly Ser Cys Cys His Gln Cys Lys Leu Leu Ala Pro Gly Thr Leu Cys 10 Arg Glu Gln Ala Arg Gln Cys Asp Leu Pro Glu Phe Cys Thr Gly Lys Ser Pro His Cys Pro Thr Asn Phe Tyr Gln Met Asp Gly Thr Pro Cys 65 70 75 15 Glu Gly Gly Gln Ala Tyr Cys Tyr Asn Gly Met Cys Leu Thr Tyr Gln 85 90 Glu Gln Cys Gln Gln Leu Trp Gly Pro Gly Ala Arg Pro Ala Pro Asp 105 110 Leu Cys Phe Glu Lys Val Asn Val Ala Gly Asp Thr Phe Gly Asn Cys 20 115 120 125 Gly Lys Asp 130
- 25 SEQ ID NO:15:

30

35

40

45

50

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1183 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: not relevant
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

					CT G hr G												46
5																TGT	94
	Asn	Arg	Arg	Glu	Leu	Авр	Arg	Tyr	Leu	Gln	Ser	Gly	Gly	Gly	Met	Cys	
	CTC	TCC	AAC	ATG	CCA	GAC	ACC	AGG	ATG	TTG	TAT	GGA	GGC	CGG	AGG	TGT	142
10	Leu	Ser	Asn	Met	Pro	Asp	Thr	Arg	Met	Leu	Tyr	Gly	Gly	Arg	Arg	Сув	
,,	GGG	AAC	GGG	TAT	CTG	GAA	GAT	GGG	GAA	GAG	TGT	GAC	TGT	GGA	GAA	GAA	190
	Gly	Asn	Gly	Tyr	Leu	Glu	Asp	Gly	Glu	Glu	Сув	qaA	Сув	Gly	Glu	Glu	
	GAG	GAA	TGT	AAC	AAC	CCC	TGC	TGC	AAT	GCC	TCT	AAT	TGT	ACC	CTG	AGG	238
15	Glu	Glu	Сув	Asn	Asn	Pro	Cys	Cys	Asn	Ala	Ser	Asn	Cys	Thr	Leu	Arg	
	CCG	GGG	GCG	GAG	TGT	GCT	CAC	GGC	TCC	TGC	TGC	CAC	CAG	TGT	AAG	CTG	286
	Pro	Gly	Ala	Glu	Cys	Ala	His	Gly	Ser	Сув	Сув	His	Gln	Cys	Lys	Leu	
20	TTG	GCT	CCT	GGG	ACC	CTG	TGC	CGC	GAG	CAG	GCC	AGG	CAG	TGT	GAC	CTC	334
	Leu	Ala	Pro	Gly	Thr	Leu	Сув	Arg	Glu	Gln	Ala	Arg	Gln	Сув	Asp	Leu	
	CCG	GAG	TTC	TGT	ACG	GGC	AAG	TCT	CCC	CAC	TGC	ССТ	ACC	AAC	TTC	TAC	382
25	Pro	Glu	Phe	Cys	Thr	Gly	Lys	Ser	Pro	His	Сув	Pro	Thr	Asn	Phe	Tyr	

5

	CNG	DTG.	ТАЭ	ССТ	ACC	CCC	TGT	GAG	GGC	GGC	CAG	GCC	TAC	TGC	TAC	AAC	430
								Glu									
5								GAG									478
	Gly	Met	Cys	Leu	Thr	Tyr	GIN	Glu	GIN	Сув	Gin	GIN	ren	Trp	GIA	Pro	
	GGA	GCC	CGA	CCT	GCC	CCT	GAC	CTC	TGC	TTC	GAG	AAG	GTG	AAT	GTG	GCA	526
								Leu									
10																	
								GGA									574
	GIA	Asp	Thr	Pne	GIÀ	ASD	Сув	Gly	гув	Asp	met	ABN	GIA	GIU	415	Arg	
	AAG	TGC	AAC	ATG	AGA	GAT	GCG	AAG	TGT	GGG	AAG	ATC	CAG	TGT	CAG	AGC	622
15								Lys									
														<b></b>			670
								TCC Ser									670
	361	914	ALG	æy			024	502									
20								ATC									718
	Ile	Ile	Met	Asn	Gly	Arg	Gln	Ile	Gln	Сув	Arg	Gly	Thr	His	Val	Tyr	
	~~»	CCM	com	CAC	CNC	CNC	<b>ሪርጥ</b>	GAC	ATC	CTG	GAC	CCA	ccc	CTG	GTG	ATYC	766
								Asp									,,,,
25	_	_															
								AAC									814
	Thr	Gly	Thr	Lys	Cys	Gly	Tyr	neA	His	Ile	Сув	Leu	Glu	Gly	Gln	Сув	
	AGG	AAC	ACC	TCC	TTC	TTT	GAA	ACT	GAA	GGC	TGT	GGG	AAG	AAG	TGC	AAT	862
30								Thr									
								AAC Asn									910
	GIY	HIB	GIY	vai	Cys	ASD	WBII	KSII	GIII	ABU	Cys	110	Cys	Dea	FIU	Gry	
35								ACA									958
	Trp	Ala	Pro	Pro	Phe	Cys	Asn	Thr	Pro	Gly	His	Gly	Gly	Ser	Ile	Asp	
	N.C.T	cca	COT	איניאל	ccc	ርርሞ	GAG	AGT	CTC	CCT	CCT	GTG	GTA	GCT	GGA	GTG	1006
								Ser									
40		_															
								GCG									1054
	Leu	Val	Ala	He	Leu	vaı	Leu	Ala	vaı	Leu	met	Leu	met	TYE	Tyr	Cys	
	TGC	AGA	CAG	AAC	AAC	AAA	CTA	GGC	CAA	CTC	AAG	CCC	TCA	GCT	CTC	CCT	1102
45								Gly									
	mc				<b></b>	C 2 C	mm^	nc#	- Alfredia	ccc	4	200	Citate	طحاط	CDG	ממ	1150
								AGT Ser									1130
		-,0		3					-, -			3					
50								CCA									1183
	Ser	Gly	Thr	Gly	His	Ala	Asn	Pro	Thr	Phe	Lys						

SEQ ID NO:16:

55

# (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 394 amino acids

- (B) TYPE: amino acid (D) TOPOLOGY: linear

5

(ii) MOLECULE TYPE: protein (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

	Gly Al	a Ala	Thr	Gly 5	His	Pro	Phe	Pro	Lys 10	Val	Phe	Asn	Gly	Сув 15	Asn
10	Arg Ar	g Glu	Leu 20	qeA	Arg	Tyr	Leu	Gln 25	Ser	Gly	Gly	Gly	Met 30	Сув	Leu
	Ser As	n Met 35		Asp	Thr	Arg	Met 40	Leu	Tyr	Gly	Gly	Arg 45	Arg	Сув	Gly
15	Asn Gl	y Tyr O	Leu	Glu	Авр	Gly 55	Glu	Glu	Cys	qaA	Сув 60	Gly	Glu	Glu	Glu
	Glu Cy 65	neA e	Asn	Pro	Сув 70	Сув	aeA	Ala	Ser	Asn 75	Cys	Thr	Leu	Arg	Pro 80
	Gly Al		•	85		_		-	90					95	
20	Ala Pr		100					105					110		
	Glu Ph	115		_	_		120					125			
25	Met As 13 Met Cy	0			_	135					140				
	145			_	150					155					160
	Ala Ar	g Pro	Ala	Pro 165	Asp	Leu	Сув	Phe	Glu 170	Lys	Val	Asn	Val	Ala 175	Gly
30	Asp Th		180					185					190		
	Cys As	195					200					205			
35	Glu Al 21	0				215					220	_			
	Ile Me	t Asn	Gly	Arg	Gln 230	Ile	Gln	Суѕ	Arg	Gly 235	Thr	His	Val	Tyr	Arg 240
	Gly Pr			245		_			250					255	
40	Gly Th	_	260		_			265					270		
	Asn Th	275					280	_			_	285			
45	His Gl	0	_			295					300				
	Ala Pr 305			_	310					315					320
	Gly Pr			325					330					335	
50	Val Al		340					345					350		
	Arg Gl	355					360					365			
55	Lys Le	_	Gln	Gln	Phe	Ser 375	Сув	Pro	Phe	Arg	Val 380	Ser	Gln	Asn	Ser
	Gly Th		His	Ala	Asn 390		Thr	Phe	Lys						

5

	SEQ ID NO:17:
	(i) SEQUENCE CHARACTERISTICS:
5	<ul><li>(A) LENGTH: 624 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: not relevant</li><li>(D) TOPOLOGY: linear</li></ul>
10	(ii) MOLECULE TYPE: cDNA (iii) HYPOTHETICAL: NO (iv) ANTI-SENSE: NO (vii) IMMEDIATE SOURCE:
15	(B) CLONE:CLONE TM
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:
20	
25	
30	

	GC															CAA	47
		Thr	гув	Сув	Ala	ASP	GTÅ	гуя	116	Сув	Leu	MBII	Aig	GIII	Cys	Gln	
5																GGC	95
	Ası	ılle	Sez	r Val	Phe	: Gly	, Val	. His	3 Glu	Сув	Ala	Met	Gln	Cys	His	Gly	
	AGZ	A GGG	GTC	3 TGC	: AAC	: AAC	: AGG	AAG	AAC	TGC	CAC	TGC	GAG	GCC	CAC	TGG	143
	Arg	g Gly	/ Val	L Cys	a Aen	Asn	Arg	Lys	aaA s	Cys	His	Cys	Glu	Ala	Hie	Trp	
10																	
																AGC	191
	Ala	a Pro	Pro	) Phe	e Cys	asp	ь гля	Pne	e Gry	Pne	GIY	GI	Sei	The	ASL	Ser	
	GGG	ccc	ATC	CGG	CAR	GCA	GAT	· AAC	CAA	GGI	TTA	ACC	: ATA	GGA	ATI	CTG	239
15																Leu	
										· cci	- Turbut	- C374	· (274	* ***	CTC	AAA :	287
																Lys	20,
	V &I.		. 114	. Dec	. Cys	<u> </u>	. Dec		. ,	. Oly	- 110	, ,,,		-7-		. –,,–	
20																TTA:	335
	Arg	, Lys	Thi	r Leu	Ile	Arg	Leu	Lei	ı Phe	Thr	Asn	Lys	Lys	Thr	Thr	: Ile	
	GAZ		. CTI	A AGG	ייבאר ב	GTC	: CGC	ce	TCC	. ccc	CCA	ccc	CGT	GGC	TTC	CAA	383
																Gln	
25		_											,				
																AAG	431
	Pro	Cys	Glr	ı Ala	His	Leu	Gly	Hie	Leu	Gly	rya	Gly	Leu	Met	Arg	Lys	
	ccc	2 CC2	GA7	ቦ ጥሮር	י דאר	· cca	CCG	. AAC	GAC	: AAT	. ccc	AGG	AGA	TTG	CTG	CAG	479
																Gln	
30					_												
																CCT	527
	Cys	Glr	Ası	ı Val	l Asp	Ile	Ser	Arg	Pro	Leu	Asn	Gly	Leu	Asn	Val	Pro	
	CAC	3 CCC	CAC	G TCF	A ACT	CAG	CGA	GTY	CTI	CCI	ccc	CTC	: CAC	: CGG	GCT	CCA	575
35																Pro	
	CG	r GCI	CC	r ago	GTC	CCT	GCC	: AG	CCC	CTG	CCA	GCC	AAG	CCT	GC	CTT	623
	Arg	, Ala	Pro	Ser	· Val	Pro	Ala	Arg	Pro	Lev	Pro	Ala	Lys	Pro	Ala	Leu	
40	A																624
	~																3-

SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 207 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

50

45

- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

	Thr 1	Lys	Сув	Ala	Yab 2	Gly	Lys	Ile	Сув	Leu 10	Asn	Arg	Gln	Суз	Gln 15	Asn
5	Ile	Ser	Val	Phe 20	Gly	Val	His	Glu	Сув 25	Ala	Met	Gln	Сув	His 30	Gly	Arg
	Gly	Val	Сув 35	Asn	Asn	Arg	Lys	Asn 40	Сув	His	Сув	Glu	Ala 45	His	Trp	Ala
	Pro	Pro 50	Phe	Сув	Aap	ГÀа	Phe 55	Gly	Phe	Gly	Gly	Ser 60	Thr	qaA	Ser	Gly
10	Pro 65	Ile	Arg	Gln	Ala	Asp 70	Asn	Gln	Gly	Leu	Thr 75	Ile	Gly	Ile	Leu	Val 80
	Thr	Ile	Leu	Сув	Leu 85	Leu	Ala	Ala	Gly	Phe 90	Val	Val	Tyr	Leu	Lys 95	Arg
15	Lys	Thr	Leu	Ile 100	Arg	Leu	Leu	Phe	Thr 105	Asn	Lys	Lys	Thr	Thr 110	Ile	Glu
•	Lys	Leu	Arg 115	Cys	Val	Arg	Pro	Ser 120	Arg	Pro	Pro	Arg	Gly 125	Phe	Gln	Pro
	Суз	Gln 130	Ala	His	Leu	Gly	His 135	Leu	Gly	Lys	Gly	Leu 140	Met	Arg	Lys	Pro
20	Pro 145	Asp	Ser	Tyr	Pro	Pro 150	Lys	Asp	Asn	Pro	Arg 155	Arg	Leu	Leu	Gln	Cys 160
	Gln	Asn	Val	Asp	Ile 165	Ser	Arg	Pro	Leu	Asn 170	Gly	Leu	Asn	Val	Pro 175	Gln
25	Pro	Gln	Ser	Thr 180	Gln	Arg	Val	Leu	Pro 185	Pro	Leu	His	Arg	Ala 190	Pro	Arg
	Ala	Pro	Ser 195	Val	Pro	Ala	Arg	Pro 200	Leu	Pro	Ala	Lys	Pro 205	Ala	Leu	

#### 30 SEQ ID NO:19:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2669 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vii) IMMEDIATE SOURCE
  - (B) CLONE:JM109 (pMel  $\beta$ -24C) JM109 (pMel  $\beta$ -24N)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:
- 50 C GGA GCT GCC ACT GGG CAC CCC TTT CCC AAA GTG TTC AAT GGA TGC
  Gly Ala Ala Thr Gly His Pro Phe Pro Lys Val Phe Asn Gly Cys
  - AAC AGG AGG GAG CTG GAC AGG TAT CTG CAG TCA GGT GGT GGA ATG TGT 94

55

35

40

	Asn	Arg	Arg	Glu	Leu	Asp	Arg	Tyr	Leu	Gln	Ser	Gly	Gly	Gly	Met	Cys	
	C-TC-	TCC	AAC	ATG	CCA	GAC	ACC	AGG	ATG	TTG	TAT	GGA	GGC	CGG	AGG	TGT	142
5																Сув	
																GAA Glu	190
	GTÅ	ASII	GIY	ıyı	Deu	GIU	wap	Gry	GIU	GIU	Cys	veh	Cys	GLY	GIU	<b>3.</b> u	
10	GAG	GAA	TGT	AAC	AAC	CCC	TGC	TGC	AAT	GCC	TCT	AAT	TGT	ACC	CTG	AGG	238
	Glu	Glu	Сув	Asn	Asn	Pro	Сув	CAa	Asn	Ala	Ser	Asn	Сув	Thr	Leu	Arg	
	~~~	000	~~~	CNG	TYOT	CCT	CAC	ccc	TOO	TYCE	TYCEC	CAC	CAG	тст	AAG	CTG	286
																Leu	200
15		•			_			_		_							
																CTC	334
	Leu	Ala	Pro	Gly	Thr	Leu	Cys	Arg	GIU	GIN	Ala	arg	GIN	Сув	Авр	Leu	
	CCG	GAG	TTC	TGT	ACG	GGC	AAG	TCT	CCC	CAC	TGC	CCT	ACC	AAC	TTC	TAC	382
20									Pro								
							-	~~~	~~~	~~~	~~~		ma.c	mcc.	The Co	220	430
									Gly							AAC Asn	430
	9211		,,	0-,			-7-		7	3			- , -		•		
25																CCC	478
	GJA	Met	Cys	Leu	Thr	Tyr	Gln	Glu	Gln	Сув	Gln	Gln	Leu	Trp	Gly	Pro	
	GGA	acc	CGA	CCT	GCC	CCT	GAC	CTC	TGC	TTC	GAG	AAG	GTG	AAT	GTG	GCA	526
																Ala	
30																	
																AGG Arg	574
	GIA	A <sub>C</sub>	****	7110	OL,	70.11	-,5	O.J	_,_	, LUP			<b>-</b> -,			3	
									TGT								622
35	Lys	Сув	Asn	Met	Arg	qaA	Ala	Lys	Сув	Gly	Lys	Ile	Gln	Сув	Gln	Ser	
	тст	GAG	GCC	CGG	CCC	CTG	GAG	TCC	AAC	GCG	GTG	ccc	ATT	GAC	ACC	ACT	670
									Asn								
														<b>~</b>	000	mag	710
40																TAC Tyr	718
	CGA	GGT	CCT	GAG	GAG	GAG	GGT	GAC	ATG	CTG	GAC	CCA	GGG	CTG	GTG	ATG	766
	Arg	Gly	Pro	Glu	Glu	Glu	Gly	Asp	Met	Leu	Asp	Pro	GIA	Leu	Val	Met	
45	ACT	GGA	ACC	AAG	TGT	GGC	TAC	AAC	CAT	ATT	TGC	CTT	GAG	GGG	CAG	TGC	814
	Thr	Gly	Thr	Lys	Суз	Gly	Tyr	Asn	His	Ile	Cys	Leu	Glu	Gly	Gln	Сув	
									<b></b>	~~~		~~~			<b>T</b>	220	963
									GAA Glu								862
50	~~9			JUL	- 40					;	-1-		_3_	-10	2		
									CAG								910
	Gly	His	Gly	Val	Сув	Asn	neA	Asn	Gln	Asn	Сув	His	Сув	Leu	Pro	Gly	
	TCG	GCC	CCG	CCC	TTC	TGC	AAC	ACA	CCG	GGC	CAC	GGG	GGC	AGT	ATC	GAC	958
55									Pro								

	AGT GGG CCT ATG CCC CCT GAG AGT GTG GGT CCT GTG GTA GCT GGA GTG Ser Gly Pro Met Pro Pro Glu Ser Val Gly Pro Val Val Ala Gly Val	1006
5	TTG GTG GCC ATC TTG GTG CTG GCG GTC CTC ATG CTG ATG TAC TAC TGC Leu Val Ala Ile Leu Val Leu Ala Val Leu Met Leu Met Tyr Tyr Cys	1054
	TGC AGA CAG AAC AAA CTA GGC CAA CTC AAG CCC TCA GCT CTC CCT Cys Arg Gln Asn Asn Lys Leu Gly Gln Leu Lys Pro Ser Ala Leu Pro	1102
10	TCC AAG CTG AGG CAA CAG TTC AGT TGT CCC TTC AGG GTT TCT CAG AAC	1150
15	Ser Lys Leu Arg Gln Gln Phe Ser Cys Pro Phe Arg Val Ser Gln Asn  AGC GGG ACT GGT CAT GCC AAC CCA ACT TTC AAG CCG GAA TTC CGG GCC	1198
	Ser Gly Thr Gly His Ala Asn Pro Thr Phe Lys Pro Glu Phe Arg Ala  CCC CAC AGC CCA CAC CAC CAT GAC AAG GGC CAC CAA TTC CAC GGC CAC	1246
20	Pro His Ser Pro His His His Asp Lys Gly His Gln Phe His Gly His  ACC CTC CTC CAC TCT GGG GAC GAC CCG GAT CCT CAC TGA GCTGACCACA	1295
	Thr Leu Leu His Ser Gly Asp Asp Pro Asp Pro His *  ACAGCCACTA CAACTGCAGC CACTGGATCC ACGGCCACCC TGTCCTCCAC CCCAGGGACC	1355
25	ACCTGGATCC TCACAGAGCC GAGCACTATA GCCACCGTGA TGGTGCCCAC CGGTTCCACG GCCACCGCCT CCTCCACTCT GGGAACAGCT CACACCCCCA AAGTGGTGAC CACCATGGCC ACTATGCCCA CAGCCACTGC CTCCACGGTT CCCAGCTCGT CCACCGTGGG GACCACCCGC	1415 1475 1535
	ACCCCTGCAG TGCTCCCCAG CAGCCTGCCA ACCTTCAGCG TGTCCACTGT GTCCTCCTCA GTCCTCACCA CCCTGAGACC CACTGGCTTC CCCAGCTCCC ACTTCTCTAC TCCCTGCTTC TGCAGGGCAT TTGGACAGTT TTTCTCGCCC GGGGAAGTCA TCTACAATAA GACCGACCGA	1595 1655 1715
30	GCCGGCTGCC ATTTCTACGC AGTGTGCAAT CAGCACTGTG ACATTGACCG CTTCCAGGGC GCCTGTCCCA CCTCCCCACC GCCAGTGTCC TCCGCCCGC TGTCCTCGCC CTCCCCTGCC CCTGGCTGTG ACAATGCCAT CCCTCTCCGG CAGGTGAATG AGACCTGGAC CCTGGAGAAC	1775 1835 1895
35	TGCACGGTGG CCAGGTGCGT GGGTGACAAC CGTGTCGTCC TGCTGGACCC AAAGCCTGTG GCCAACGTCA CCTGCGTGAA CAAGCACCTG CCCATCAAAG TGTCGGACCC GAGCCAGCCC TGTGACTTCC ACTATGAGTG CGAGTGCATC TGCAGCATGT GGGGCGGCTC CCACTATTCC	1955 2015 2075
	ACCTTTGACG GCACCTCTTA CACCTTCCGG GGCAACTGCA CCTATGTCCT CATGAGAGAG ATCCATGCAC GCTTTGGGAA TCTCAGCCTC TACCTGGACA ACCACTACTG CACGGCCTCT GCCACTGCCG CTGCCGCCCG CTGCCCCCGC GCCCTCAGCA TCCACTACAA GTCCATGGAT	2135 2195 2255
40	ATCGTCCTCA CTGTCACCAT GGTGCATGGG AAGGAGGAGG GCCTGATCCT GTTTGACCAA ATTCCGGTGA GCAGCGGTTT CAGCAAGAAC GGCGTGCTTG TGTCTGTGCT GGGGACCACC ACCATGCGTG TGGACATTCC TGCCCTGGGC GTGAGCGTCA CCTTCAATGG CCAAGTCTTC	2315 2375 2435
	CAGGCCCGGC TGCCCTACAG CCTCTTCCAC AACAACACCG AGGGCCAGTG CGGCACCTGC ACCAACAACC AGAGGGACGA CTGTCTCCAG CGGGACGGAA CCACTGCCGC CAGTTGCAAG GACATGGCCA AGACGTGGCT GGTCCCCGAC AGCAGAAAGG ATGGCTGCTG GGCCCCGACT	2495 2495 2555 2615
45	GGCACACCCC CCACTGCCAG CCCCGCAGCC CCGGTGTCTA GCACACCCAC CCCG	2669

### SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 427 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MC

50

55

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

	1				5					10					Cys 15	
5	Arg	Arg	Glu	Leu 20	qaA	Arg	Tyr	Leu	Gln 25	Ser	Gly	Gly	Gly	Met 30	Сув	Leu
	Ser	Asn	Met 35	Pro	Asp	Thr	Arg	Met 40	Leu	Tyr	Gly	Gly	Arg 45	Arg	Cys	Gly
	Asn	Gly 50	Tyr	Leu	Glu	Asp	Gly 55	Glu	Glu	Сув	qeA	Сув 60	Gly	Glu	Glu	Glu
10	Glu 65	Cys	Asn	Asn	Pro	Cys 70	Сув	Asn	Ala	Ser	Asn 75	Сув	Thr	Leu	Arg	Pro 80
				Суз	85					90					95	
15			_	Thr 100		_			105					110		
			115	Thr				120					125			
20		130	-	Thr			135					140				
20	145	-		Thr	_	150					155					160
		_		Ala	165					170					175	
25	_			Gly 180 Arg					185					190		
	•		195	Pro	_			200					205			
30		210	_	Gly			215					220				
	225			Glu		230					235					240
	_			Сув	245					250					255	
35	•		_	260 Phe	_				265					270		
			275	Суз				280					285			
40	Ala	290 Pro	Pro	Phe	Сув	Asn	295 Thr	Pro	Gly	His	Gly	300 Gly	Ser	Ile	Asp	
	305 Gly	Pro	Met	Pro	Pro	310 Glu	Ser	Val	Gly	Pro	315 Val	Val	Ala	Gly	Val	320 Leu
45	Val	Ala	Ile	Leu	325 Val	Leu	Ala	Val		330 Met	Leu	Met	Tyr		Cys	Сув
45	Arg	Gln		Asn	Lys	Leu	Gly		345 Leu	Lys	Pro	Ser	Ala 365	350 Leu	Pro	Ser
	Lys		355 Arg	Gln	Gln	Phe	Ser 375	360 Cys	Pro	Phe	Arg	Val 380		Gln	Asn	Ser
50	Gly 385	370 Thr	Gly	His	Ala	Asn 390	-	Thr	Phe	Lys	Pro 395		Phe	Arg	Ala	Pro
		Ser	Pro	His	His 405		Asp	Lys	Gly	His 410		Phe	His	Gly	His 415	
55	Leu	Leu	His	Ser 420		Двр	qeA	Pro	Asp 425		His					

SEQ ID N0:21:

	(i) SEQUENCE CHARACTERISTICS:
5	<ul><li>(A) LENGTH: 1483 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: not relevant</li><li>(D) TOPOLOGY: linear</li></ul>
10	(ii) MOLECULE TYPE: cDNA (iii) HYPOTHETICAL: NO (iv) ANTI-SENSE: NO (vii) IMMEDIATE SOURCE
	(B) CLONE:JM109 (pMel α-25C)
15	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:
20	
25	
<i>3</i> 0	
<i>3</i> 5	
40	
45	
. <del>-</del>	
50	

	GAT	GGG	CAC	TCA	TGT	CAG	GAT	GTG	GAC	GGC	TAC	TGC	TAC	AAT	GGC	ATC	48
	Asp	Gly	His	Ser	Сув	Gln	Asp	Val	Asp	Gly	Tyr	Cys	Tyr	Asn	Gly	Ile	
5			200			380	<b></b>		<b></b>		-						
																GAT	144
	rys	Pro	ATA	Pro	GIĄ	IIe	Сув	Pne	GIU	Arg	Val	Asn	Ser	Ala	Gly	Asp	
	CCT	TAT	GGC	AAC	TGT	GGC	AAA	GTC	TCG	AAG	AGT	TCC	TTT	GCC	AAA	TGC	192
10	Pro	Tyr	Gly	Asn	Cys	Gly	Lys	Val	Ser	Lys	Ser	Ser	Phe	Ala	Lys	Сув	
	GAG	ATG	AGA	GAT	GCT	AAA	TGT	GGA	AAA	ATC	CAG	TGT	CAA	GGA	GGT	GCC	240
	Glu	Met	Arg	Авр	Ala	Lys	Сув	Gly	Lys	Ile	Gln	Сув	Gln	Gly	Gly	Ala	
	200	~~~	CCA	CTC.	እጥጥ	CCT	N.C.C	B 200	ccc	- COTON	maa	A TO B	<b>C.3.3</b>	101	220	ATC	200
15																Ile	288
	Ser	Arg	PIO	Val	116	GIY	IHI	MBII	ATA	vai	ser	IIe	GIU	Tur	ABN	iie	
																TAC	336
	Pro	Leu	Gln	Gln	Gly	Gly	Arg	Ile	Leu	Cys	Arg	Gly	Thr	His	Val	Tyr	
20	TTG	GGC	GAT	GAC	ATG	CCG	GAC	CCA	GGG	CTT	GTG	CTT	GCA	GGC	ACA	AAG	384
	Leu	Gly	Asp	Asp	Met	Pro	qaA	Pro	Gly	Leu	Val	Leu	Ala	Gly	Thr	Lys	
	ጥርጥ	GCA	GAT	GGA	AAA	ልጥሮ	TGC	CTC	ልልጥ	CCT	CAA	TCT	CAA	AAT	ል የ	እርሞ	432
					Lys												132
25	C,U			017	<b></b> ,		-,-			,	<b>021.</b>	-,u	02			001	
	GTC	TTT	GGG	GTT	CAC	GAG	TGT	GCA	ATG	CAG	TGC	CAC	GGC	AGA	GGG	GTG	480
	Val	Phe	Gly	Val	His	Glu	Cys	Ala	Met	Gln	Сув	His	Gly	Arg	Gly	Val	
	TGC	AAC	AAC	AGG	AAG	AAC	TGC	CAC	TGC	GAG	GCC	CAC	TGG	GCA	ССТ	ccc	528
30	Сув	Asn	Asn	Arg	Lys	Asn	Сув	His	Cys	Glu	Ala	His	Trp	Ala	Pro	Pro	
	mm/C	m <nm< td=""><td>CNO</td><td>220</td><td>TTT</td><td>ccc</td><td>deter</td><td>CCA</td><td>CCA</td><td>n.c.c</td><td>N.C.N</td><td>CNO</td><td>200</td><td>ccc</td><td>CCC</td><td>200</td><td>576</td></nm<>	CNO	220	TTT	ccc	deter	CCA	CCA	n.c.c	N.C.N	CNO	200	ccc	CCC	200	576
																	5/6
	Pile	Cys	wab	гур	Phe	Gry	PILE	GIY	GIY	Ser	Int	мар	SeI	GIY	PIO	116	
35	CGG	CAA	GCA	GAA	GCA	AGG	CAG	GAA	GCT	GCA	GAG	TCC	AAC	AGG	GAG	CGC	624
	Arg	Gln	Ala	Glu	Ala	Arg	Gln	Glu	Ala	Ala	Glu	Ser	Asn	Arg	Glu	Arg	
	GGC	CNG	ccc	CNG	GAG	ccc	OTC:	CCN	TCG	CNG	CNC	CAT	ccc	тст	A (~T	ccc	672
					_												0/2
40	GIÀ	GIII	GTÅ	GTII	Glu	FLU	ACTT	GIY	361	GIH	GIU	urg	wra	SET	1117	wrg	
	TCA	CTG	ACA	CTC	ATC	TGA	GCCC	TCCC	AT C	ACAT	GGAG	A CO	GTGA	CCAC	;		720
	Ser	Leu	Thr	Leu	Ile	*											

	TGCTGCTGCA	GAGGAGGTCA	CGCGTCCCCA	AGGCCTCCTG	TGACTGGCAG	CATTGACTCT	780
	GTGGCTTTGC	CATCGTTTCC	ATGACAACAG	ACACAACACA	GTTCTCGGGG	CTCAGGAGGG	840
5	GAAGTCCAGC	CTACCAGGCA	CGTCTGCAGA	AACAGTGCAA	GGAAGGGCAG	CGACTTCCTG	900
5	GTTGAGCTTC	TGCTAAAACA	TGGACATGCT	TCAGTGCTGC	TCCTGAGAGA	GTAGCAGGTT	960
	ACCACTCTGG	CAGGCCCCAG	CCCTGCAGCA	AGGAGGAAGA	GGACTCAAAA	GTCTGGCCTT	1020
	TCACTGAGCC	CCCACAGCAG	TGGGGGAGAA	GCAAGGGTTG	GGCCCAGTGT	CCCCTTTCCC	1080
	CAGTGACACC	TCAGCCTTGG	CAGCCCTGAT	GACTGGTCTC	TGGCTGCAAC	TTAATGCTCT	1140
10	GATATGGCTT	TTAGCATTTA	TTATATGAAA	ATAGCAGGGT	TTTAGTTTTT	AATTTATCAG	1200
	AGACCCTGCC	ACCCATTCCA	TCTCCATCCA	AGCAAACTGA	ATGGCATTGA	AACAAACTGG	1260
	AGAAGAAGGT	AGGAGAAAGG	GCGGTGAACT	CTGGCTCTTT	GCTGTGGACA	TGCGTGACCA	1320
	GCAGTACTCA	GGTTTGAGGG	TTTGCAGAAA	GCCAGGGAAC	CCACAGAGTC	ACCAACCCTT	1380
	CATTTAACAA	GTAAGAATGT	TAAAAAGTGA	AAACAATGTA	AGAGCCTAAC	TCCATCCCCC	1440
15	GTGGCCATTA	CTGCATAAAA	TAGAGTGCAT	CCCGCCCGAA	TTC		1483

# (2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 230 amino acids

(B) TYPE: amino acid
(D) TOPOLOGY: linear

25

20

30

35

40

45

50

55

(ii) MOLECULE TYPE: protein (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

Asp Gly His Ser Cys Gln Asp Val Asp Gly Tyr Cys Tyr Asn Gly Ile 10 Cys Gln Thr His Glu Gln Gln Cys Val Thr Leu Trp Gly Pro Gly Ala 5 Lys Pro Ala Pro Gly Ile Cys Phe Glu Arg Val Asn Ser Ala Gly Asp Pro Tyr Gly Asn Cys Gly Lys Val Ser Lys Ser Ser Phe Ala Lys Cys 10 Glu Met Arg Asp Ala Lys Cys Gly Lys Ile Gln Cys Gln Gly Gly Ala 70 75 Ser Arg Pro Val Ile Gly Thr Asn Ala Val Ser Ile Glu Thr Asn Ile 90 Pro Leu Gln Gln Gly Gly Arg Ile Leu Cys Arg Gly Thr His Val Tyr 15 100 105 Leu Gly Asp Asp Met Pro Asp Pro Gly Leu Val Leu Ala Gly Thr Lys 120 Cys Ala Asp Gly Lys Ile Cys Leu Asn Arg Gln Cys Gln Asn Ile Ser 135 140 Val Phe Gly Val His Glu Cys Ala Met Gln Cys His Gly Arg Gly Val 20 155 150 Cys Asn Asn Arg Lys Asn Cys His Cys Glu Ala His Trp Ala Pro Pro 165 170 Phe Cys Asp Lys Phe Gly Phe Gly Gly Ser Thr Asp Ser Gly Pro Ile 25 180 185 Arg Gln Ala Glu Ala Arg Gln Glu Ala Ala Glu Ser Asn Arg Glu Arg 200 Gly Gln Gly Gln Glu Pro Val Gly Ser Gln Glu His Ala Ser Thr Ala 215 220 Ser Leu Thr Leu Ile \* 30 225 230

### SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1569 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vii) IMMEDIATE SOURCE
  - (B) CLONE:JM109 (pMel α-26N)
- 50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

55

35

40

	GGG	GAC	CTC	TGG	ATC	CCA	GTG	AAG	AGC	TTC	GAC	TCC	AAG	AAT	CAT	CCA	48
	Gly	Asp	Leu	Trp	Ile	Pro	Val	Lys	Ser	Phe	Asp	Ser	Lys	Asn	His	Pro	
5	GNA	GTYG	CTC	<b>ד</b> ממ	ልሞኮ	CGA	CTP A	CAA	CGG	GAA	AGC	AAA	GAA	CTG	ATC	ATA	96
•																Ile	
									_			_					
									ATT								144
10	Asn	Leu	Glu	Arg	Asn	Glu	Gly	Leu	Ile	Ala	Ser	Ser	Phe	Thr	GIU	Thr	
10	CAC	тат	CTG	CAA	GAC	GGT	ACT	GAT	GTC	TCC	CTC	GCT	CGA	AAT	TAC	ACG	192
																Thr	
		-															
																GCA	240
15	Gly	His	Сув	Tyr	Tyr	H18	GIÀ	H18	vai	Arg	GIA	TYT	ser	Asp	ser	Ala	
	GTC	AGT	CTC	AGC	ACG	TGT	TCT	GGT	CTC	AGG	GGA	CTT	ATT	GGG	TTT	GAA	288
																Glu	
20									ATG							TAC Tyr	336
	ABN	GIU	ser	туг	vaı	Leu	GIU	PIG	Met	nya	261	Ala	1111	ASII	AL Y	.,.	
	AAA	CTC	TTC	CCA	GCG	AAG	AAG	CTG	AAA	AGC	GTC	CGG	GGA	TCA	TGT	GGA	384
	Lys	Leu	Phe	Pro	Ala	Lys	Lys	Leu	Lys	Ser	Val	Arg	Gly	Ser	Сув	Gly	
25						663		~m~				3 B 777	CENC	- Children	~~~	CCA	432
																CCA Pro	732
	361	nis	UTD	<b>NO</b> II	1114		2000-		,,,,,	<i>,</i>	2,5						
																GCA	480
30	Pro	Ser	Gln	Thr	Trp	Ala	Arg	Arg	His	Lys	Arg	Glu	Thr	Leu	Lys	Ala	
				a <b>m</b> a	<b>636</b>	~~~	CTC.	N TO C	CTC	CCA	CNC	220	CCA	GNG	ملحلص	CAG	528
																Gln	320
		275	-1-														
35																ATT	576
	Arg	Gln	Gly	Lys	qaA	Leu	Glu	Lys	Val	Lys	Gln	Arg	Leu	Ile	Glu	Ile	
	COT	220	CAC	Citer	GAC	AAC	لملمك	TAC	AGA	CCA	CTG	אאר	ልጥጥ	CGG	ATC	GTG	624
																Val	-
40	•				_	_											
	TTG	GTA	GGC						GAC								672
	•				~ 1 · ·	77-7		3	3	Mak	B	T	~~		7707	COT	

	CAG	GAC	CCA	TTC	ACC	AGC	CTC	CAT	GAA	TTT	CTG	GAC	TGG	AGG	AAG	ATG	720
						Ser											-
		-															
5																GGG	768
	Lys	Leu	Leu	Pro	Arg	Lys	Ser	His	Asp	Asn	Ala	Gln	Leu	Val	Ser	Gly	
		ጥአጥ	الملتون	CAA	ccc	ACC	»CC	እጥ <u>ር</u>	GGC	ATG	GCC	CCA	אדער	ATC	AGC	ATC	816
						Thr											310
10	742	-1-		<b>U</b> 111	027				,								
	TGC	ACG	GCA	GAC	CAG	TCT	GGG	GGA	ATT	GTC	ATG	GAC	CAT	TCA	GAC	AAT	864
	Cys	Thr	Ala	Asp	Gln	Ser	Gly	Gly	Ile	Val	Met	qaA	His	Ser	Ąsp	Asn	
15						GTG											912
	Pro	Leu	GIĄ	АТа	Ala	Val	Thr	ren	ALA	uis	GIU	Leu	GIA	uis	ASII	Fue	
	GGG	ATG	AAT	CAT	GAC	ACA	CTG	GAC	AGG	GGC	TGT	AGC	TGT	CAA	ATG	GCG	960
						Thr											
	-																
20						TGC											1008
	Val	Glu	Lys	Gly	Gly	Сув	Ile	Met	Asn	Ala	ser	Thr	GIA	TYY	PTO	Pne	
	CCC	DTG	GTG	TTC	AGC	AGT	TGC	AGC	AGG	AAG	GAC	TTG	GAG	ACC	AGC	CTG	1056
						Ser											
25																	
						GTG											1104
	Glù	Lys	Gly	Met	Gly	Val	CAa	Leu	Phe	Asn	Leu	Pro	Glu	Val	Arg	Glu	
	TOT	TTC	ccc	GCC	CAG	AAG	тст	GGG	AAC	AGA	TTT	GTG	GAA	GAA	GGA	GAG	1152
30						Lys											
			_	_									•				
						GAG											1200
	Glu	Cys	Asp	Сув	Gly	Glu	Pro	Glu	Glu	Суз	Met	Asn	Arg	Сув	Сув	Asn	
35	GCC	ACC	ACC	TCT	ACC	CTG	AAG	CCG	GAC	GCT	GTG	TGC	GCA	CAT	GGG	CTG	1248
						Leu											
						CAG											1296
40	Суз	Сув	Glu	Asp	Cys	Gln	Leu	Lys	Pro	Ala	Gly	Thr	Ala	Сув	Arg	Asp	
	TCC	»GC	AAC	ጥርር	TCT	GAC	СТС	422	GAG	TTC	TGC	ACA	GGG	GCC	AGC	CCT	1344
						Asp											
45						GTG											1392
45	His	CAa	Pro	Ala	nsA	Val	Tyr	Leu	His	Asp	Gly	His	Ser	Cys	Gln	Asp	
	cmc	CAC	ccc	ሞእሮ	TCC	TAC	אמת	GGC	ATTC	TCC	CAG	аст	CAC	GAG	CAG	CAG	1440
						Tyr											
50	TGT	GTC	ACG	CTC	TGG	GGA	CCA	<b>GGT</b>	GCT	AAA	CCT	GCC	CCT	GGG	ATC	TGC	1488
	Сув	Val	Thr	Leu	Trp	Gly	Pro	Gly	Ala	Lys	Pro	Ala	Pro	Gly	Ile	Сув	
	opionista.	~~	202	CMC	מית ב	TCT	CCN	CCT	ርስጥ	CC	тат	CCC	ממ	ጥርጥ	GGC	AAA	1536
						Ser											
55			3								- 3 -	4			•	-	

1569

GTC TCG AAG AGT TCC TTT GCC AAA TGC GAG ATG

Val Ser Lys Ser Ser Phe Ala Lys Cys Glu Met

5	SEQ ID NO:24:
	(i) SEQUENCE CHARACTERISTICS:
10	(A) LENGTH: 523 amino acids (B) TYPE: amino acid (D) TOPOLOGY: linear
15	(ii) MOLECULE TYPE: protein (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:
20	
25	
30	
<i>35</i>	
40	
45	
50	
55	

Gly	Asp	Leu	Trp	Ile	Pro	Val	Lys	Ser	Phe	Asp	Ser	Lys	Asn	His	Pro
Glu	Val	Leu	Asn	Ile	Arg	Leu	Gln	Arg	Glu	Ser	Lys	Glu	Leu	Ile	Ile
Asn	Leu	Glu	Arg	Asn	Glu	Gly	Leu	Ile	Ala	Ser	Ser	Phe	Thr	Glu	Thr
His	Tyr	Leu	Gln	Asp	Gly	Thr	Asp	Val	Ser	Leu	Ala	Arg	Asn	Tyr	Thr
Gly	His	Сув	Tyr	Tyr	His	Gly	His	Val	Arg	Gly	Tyr	Ser	Двр	Ser	Ala
Val	Ser	Leu	Ser	Thr	Сув	Ser	Gly	Leu	Arg	Gly	Leu	Ile	Gly	Phe	Glu
Asn	Glu	Ser	Tyr	Val	Leu	Glu	Pro	Met	Lys	Ser	Ala	Thr	Asn	Arg	Tyr
Lys	Leu	Phe	Pro	Ala	Lys	Lys	Leu	Lys	Ser	Val	Arg	Gly	Ser	Сув	Gly
Ser	His	His	Asn	Thr	Pro	Asn	Leu	Ala	Ala	Lys	Asn	Val	Phe	Pro	Pro
Pro ·	Ser	Gln	Thr	Trp	Ala	Arg	Arg	His	Lys		Glu 1	Thr	Leu	Lys	Ala
Thr	Lys	Tyr	Val	Glu	Leu	Val	Ile	Val	Ala	Asp	Asn	Arg	Glu	Phe	Gln
Arg	Gln	Gly	Lys	qaA	Leu	Glu	Lys	Val	Lys	Gln	Arg	Leu	Ile	Glu	Ile
Ala	Asn	His	Val	Asp	Lys	Phe	Tyr	Arg	Pro	Leu	Asn	Ile	Arg	Ile	Val
Leu	Val	Gly	Val	Glu	Val	Trp	Asn	Asp	Met	Asp	Lys	Сув	Ser	Val	Ser
Gln	Asp	Pro	Phe	Thr	Ser	Leu	His	Glu	Phe	Leu	Asp	Trp	Arg	Lys	Met
Lys	Leu	Leu	Pro	Arg	Lys	Ser	His	Авр	Asn	Ala	Gln	Leu	Val	Ser	Gly
Val	Tyr	Phe	Gln	Gly	Thr	Thr	Ile	Gly	Met	Ala	Pro	Ile	Met	Ser	Met
Сув	Thr	Ala	Asp	Gln	Ser	Gly	Gly	Ile	Val	Met	Asp	His	Ser	Asp	Asn
Pro	Leu	Gly	Ala	Ala	Val	Thr	Leu	Ala	His	Glu	Leu	Gly	His	Asn	Phe
Gly	Met	Asn	His	Asp	Thr	Leu	Asp	Arg	Gly	Сув	Ser	Сув	Gln	Met	Ala
Val	Glu	Lys	Gly	Glv	Cvs	Ile	Met	Asn	Ala	Ser	Thr	Glv	Tvr	Pro	Phe

5

Pro Met Val Phe Ser Ser Cys Ser Arg Lys Asp Leu Glu Thr Ser Leu Glu Lys Gly Met Gly Val Cys Leu Phe Asn Leu Pro Glu Val Arg Glu Ser Phe Gly Gly Gln Lys Cys Gly Asn Arg Phe Val Glu Glu Gly Glu Glu Cys Asp Cys Gly Glu Pro Glu Glu Cys Met Asn Arg Cys Cys Asn Ala Thr Thr Cys Thr Leu Lys Pro Asp Ala Val Cys Ala His Gly Leu Cys Cys Glu Asp Cys Gln Leu Lys Pro Ala Gly Thr Ala Cys Arg Asp Ser Ser Asn Ser Cys Asp Leu Pro Glu Phe Cys Thr Gly Ala Ser Pro His Cys Pro Ala Asn Val Tyr Leu His Asp Gly His Ser Cys Gln Asp Val Asp Gly Tyr Cys Tyr Asn Gly Ile Cys Gln Thr His Glu Gln Gln Cys Val Thr Leu Trp Gly Pro Gly Ala Lys Pro Ala Pro Gly Lys Val Ser Lys Ser Ser Phe Ala Lys Cys Glu Met

#### **SEQ ID NO:25:**

5

10

15

20

25

30

35

40

45

50

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 2404 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: not relevant
  - (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: cDNA
  - (iii) HYPOTHETICAL: NO
  - (iv) ANTI-SENSE: NO
- (vii) IMMEDIATE SOURCE
  - (B) CLONE:JM109 (pMel β-24C)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

Cys Cys His Gln Cys Lys Leu Leu Ala Pro Gly Thr Leu	
	ANG TOT CCC 96
CAG GCC AGG CAG TGT GAC CTC CCG GAG TTC TGT ACG GGC	
Gln Ala Arg Gln Cys Asp Leu Pro Glu Phe Cys Thr Gly	Lys Ser Pro
CAC TGC CCT ACC AAC TTC TAC CAG ATG GAT GGT ACC CCC	TGT GAG GGC 144
His Cys Pro Thr Asn Phe Tyr Gln Met Asp Gly Thr Pro	Cys Glu Gly
0	
GGC CAG GCC TAC TGC TAC AAC GGC ATG TGC CTC ACC TAC	
Gly Gln Ala Tyr Cys Tyr Asn Gly Met Cys Leu Thr Tyr	Gln Glu Gln
TGC CAG CAG CTG TGG GGA CCC GGA GCC CGA CCT GCC CCT	GAC CTC TGC 240
5 Cys Gln Gln Leu Trp Gly Pro Gly Ala Arg Pro Ala Pro	Asp Leu Cys
TTC GAG AAG GTG AAT GTG GCA GGA GAC ACC TTT GGA AAC	TGT GGA AAG 280

	Phe	Glu	Lys	Val	Asn	Val	Ala	Gly	Asp	Thr	Phe	Gly	Asn	Сув	Gly	Lys	
	GAC	ATG	AAT	GGT	GAA	CAC	AGG	AAG	TGC	AAC	ATG	AGA	GAT	GCG	AAG	TGT	336
5																Сув	330
	GGG	AAG	ATC	CAG	TGT	CAG	AGC	TCT	GAG	GCC	CGG	CCC	CTG	GAG	TCC	AAC	384
																Asn	304
10	222		000		~~~	3.00											
70																CAG Gln	432
												•	_				
																ATG	480
15	Сув	Arg	GIÀ	Thr	HIS	Val	Tyr	Arg	GIA	Pro	Glu	Glu	Glu	GIA	qaA	Met	
	CTG	GAC	CCA	GGG	CTG	GTG	ATG	ACT	GGA	ACC	AAG	TGT	GGC	TAC	AAC	CAT	528
	Leu	qaA	Pro	Gly	Leu	Val	Met	Thr	Gly	Thr	Lys	Сув	Gly	Tyr	Asn	His	
	ATT	TGC	CTT	GAG	GGG	CAG	TGC	AGG	AAC	ACC	TCC	TTC	ماحلت	GAA	ىسى ھ	GAA	576
20																Glu	370
																CAG Gln	624
	,	٠,,,	<b></b> 1	_,_	_,_	-,-		<b></b> /		-17	<b>V</b> 42	-,5	76,11	24311	74311	UZII	
25					CTG												672
	Asn	Cys	His	Cys	Leu	Pro	Gly	Trp	Ala	Pro	Pro	Phe	Сув	Asn	Thr	Pro	
	GGC	CAC	GGG	GGC	AGT	ATC	GAC	AGT	GGG	CCT	ATG	CCC	CCT	GAG	AGT	GTG	720
	Gly	His	Gly	Gly	ser	Ile	Ąsp	Ser	Gly	Pro	Met	Pro	Pro	Glu	Ser	Val	
30	ССТ	CCT	GTG	ርጥል	GCT	GGA	CTYC	<b>ም</b> ተየረ	CTC	CCC	እ <b>ጥ</b> ር	Terres	CTYC	CTC	ccc	City	768
					Ala												700
									_								
					TAC											CAA Gln	816
35	204	1700	200		-3-	-3-	-,5	Cyo	AL 9	<b>G</b> 111	ASII	<i></i>	mys	LÆU	GLY	GIII	
					GCT											-	864
	Leu	FÀa	Pro	Ser	Ala	Leu	Pro	Ser	Lys	Leu	Arg	Gln	Gln	Phe	Ser	Сув	
40	CCC	TTC	AGG	GTT	TCT	CAG	AAC	AGC	GGG	ACT	GGT	CAT	GCC	AAC	CCA	ACT	912
40	Pro	Phe	Arg	Val	Ser	Gln	Asn	Ser	Gly	Thr	Gly	His	Ala	Asn	Pro	Thr	
	TTC	DAG	CCG	GAA	TTC	CGG	GCC	ccc	CAC	AGC	CCA	CAC	CAC	СУТ	GAC	244	960
					Phe												300
45																	
	GCC				CAC												1008
	,					,							,	p	, and p		
					GCTG	ACCA	CA A	CAGC	CACT	A CA	ACTG	CAGC	CAC	TGGA	TCC		1060
50	Двр	Pro	His	•													
	ACGG	CCAC	CC T	GTCC	TCCA	c cc	CAGG	GACC	ACC	TGGA	TCC	TCAC	AGAG	CC G	AGCA	CTATA	1120
												_	_		-	CAGCT	1180
																CGGTT	1240 1300
55																GCTTC	1360
												_	_				

	CCCAGCTCCC	ACTTCTCTAC	TCCCTGCTTC	TGCAGGGCAT	TTGGACAGTT	TTTCTCGCCC	1420
	GGGGAAGTCA	TCTACAATAA	GACCGACCGA	GCCGGCTGCC	ATTTCTACGC	AGTGTGCAAT	1480
	CAGCACTGTG	ACATTGACCG	CTTCCAGGGC	GCCTGTCCCA	CCTCCCCACC	GCCAGTGTCC	1540
5	TCCGCCCCGC	TGTCCTCGCC	CTCCCCTGCC	CCTGGCTGTG	ACAATGCCAT	CCCTCTCCGG	1600
	CAGGTGAATG	AGACCTGGAC	CCTGGAGAAC	TGCACGGTGG	CCAGGTGCGT	GGGTGACAAC	1660
	CGTGTCGTCC	TGCTGGACCC	AAAGCCTGTG	GCCAACGTCA	CCTGCGTGAA	CAAGCACCTG	1720
	CCCATCAAAG	TGTCGGACCC	GAGCCAGCCC	TGTGACTTCC	<b>ACTATGAGTG</b>	CGAGTGCATC	1780
	TGCAGCATGT	GGGGCGGCTC	CCACTATTCC	ACCTTTGACG	GCACCTCTTA	CACCTTCCGG	1840
10	GGCAACTGCA	CCTATGTCCT	CATGAGAGAG	ATCCATGCAC	GCTTTGGGAA	TCTCAGCCTC	1900
	TACCTGGACA	ACCACTACTG	CACGGCCTCT	GCCACTGCCG	CTGCCGCCCG	CTGCCCCCGC	1960
	GCCCTCAGCA	TCCACTACAA	GTCCATGGAT	<b>ATCGTCCTCA</b>	CTGTCACCAT	GGTGCATGGG	2020
	AAGGAGGAGG	GCCTGATCCT	<b>GTTTGACCAA</b>	ATTCCGGTGA	GCAGCGGTTT	CAGCAAGAAC	2080
	GGCGTGCTTG	TGTCTGTGCT	GGGGACCACC	ACCATGCGTG	TGGACATTCC	TGCCCTGGGC	2140
15	GTGAGCGTCA	CCTTCAATGG	CCAAGTCTTC	CAGGCCCGGC	TGCCCTACAG	CCTCTTCCAC	2200
	AACAACACCG	AGGGCCAGTG	CGGCACCTGC	ACCAACAACC	AGAGGGACGA	CTGTCTCCAG	2260
	CGGGACGGAA	CCACTGCCGC	CAGTTGCAAG	GACATGGCCA	AGACGTGGCT	GGTCCCCGAC	2320
	AGCAGAAAGG	ATGGCTGCTG	GGCCCCGACT	GGCACACCCC	CCACTGCCAG	CCCCGCAGCC	2380
	CCGGTGTCTA	GCACACCCAC	CCCG				2404
20							

## SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 339 amino acids (B) TYPE: amino acid (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

Cys Cys His Gln Cys Lys Leu Leu Ala Pro Gly Thr Leu Cys Arg Glu 10 Gln Ala Arg Gln Cys Asp Leu Pro Glu Phe Cys Thr Gly Lys Ser Pro 25 5 His Cys Pro Thr Asn Phe Tyr Gln Met Asp Gly Thr Pro Cys Glu Gly 40 Gly Gln Ala Tyr Cys Tyr Asn Gly Met Cys Leu Thr Tyr Gln Glu Gln 10 Cys Gln Gln Leu Trp Gly Pro Gly Ala Arg Pro Ala Pro Asp Leu Cys 75 Phe Glu Lys Val Asn Val Ala Gly Asp Thr Phe Gly Asn Cys Gly Lys 85 90 Asp Met Asn Gly Glu His Arg Lys Cys Asn Met Arg Asp Ala Lys Cys 15 105 Gly Lys Ile Gln Cys Gln Ser Ser Glu Ala Arg Pro Leu Glu Ser Asn 120 Ala Val Pro Ile Asp Thr Thr Ile Ile Met Asn Gly Arg Gln Ile Gln 135 140 Cys Arg Gly Thr His Val Tyr Arg Gly Pro Glu Glu Glu Gly Asp Met 20 150 155 Leu Asp Pro Gly Leu Val Met Thr Gly Thr Lys Cys Gly Tyr Asn His 170 165 Ile Cys Leu Glu Gly Gln Cys Arg Asn Thr Ser Phe Phe Glu Thr Glu 25 185 Gly Cys Gly Lys Lys Cys Asn Gly His Gly Val Cys Asn Asn Asn Gln 195 30 Asn Cys His Cys Leu Pro Gly Trp Ala Pro Pro Phe Cys Asn Thr Pro 215 Gly His Gly Gly Ser Ile Asp Ser Gly Pro Met Pro Pro Glu Ser Val 230 Gly Pro Val Val Ala Gly Val Leu Val Ala Ile Leu Val Leu Ala Val 245 250 Leu Met Leu Met Tyr Tyr Cys Cys Arg Gln Asn Asn Lys Leu Gly Gln 265 Leu Lys Pro Ser Ala Leu Pro Ser Lys Leu Arg Gln Gln Phe Ser Cys 40 280 Pro Phe Arg Val Ser Gln Asn Ser Gly Thr Gly His Ala Asn Pro Thr 295 300 Phe Lys Pro Glu Phe Arg Ala Pro His Ser Pro His His Asp Lys 310 315 45 Gly His Gln Phe His Gly His Thr Leu Leu His Ser Gly Asp Asp Pro 330 Asp Pro His 339 50

SEQ ID NO:27:

#### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 453 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA (iii) HYPOTHETICAL: NO

		` '	NTI-S MMEC			CE											
5		(	B) CLO	ONE :J	M109	(pMei	β-24N	I)									
		(xi) S	EQUE	NCE [	DESCI	RIPTIC	N: SE	Q ID I	NO:27:	:							
10			CT G														46
15			AGG Arg													TGT Cys	94
			AAC Asn														142
20													-	-		GAA Glu	190
			-													AGG Arg	238
25			GCG Ala														286
30			CCT Pro														334
	സ്ത	GAG	יויידירי	ጥርጥ	ACG	GGC	DAA	тст	ccc	CAC	TGC	сст	ACC	AAC	TTC	TAC	382
35																Tyr	300
																AAC Asn	430
40	-		TGC Cys					GA									453
45	INF	ORM	ATION	FOR	SEQ II	D NO:	28:										
		(i) SE	QUEN	ICE C	HARA	CTER	ISTICS	S:									
50		Ì	A) LEN B) TYF D) TO	PE: am	ino ac	id	acids										
			OLEC		-			Q ID I	NO:28:	:							

Gly Ala Ala Thr Gly His Pro Phe Pro Lys Val Phe Asn Gly Cys Asn Arg Arg Glu Leu Asp Arg Tyr Leu Gln Ser Gly Gly Met Cys Leu Ser Asn Met Pro Asp Thr Arg Met Leu Tyr Gly Gly Arg Arg Cys Gly 40 Asn Gly Tyr Leu Glu Asp Gly Glu Glu Cys Asp Cys Gly Glu Glu Glu 60 Glu Cys Asn Asn Pro Cys Cys Asn Ala Ser Asn Cys Thr Leu Arg Pro 70 75 Gly Ala Glu Cys Ala His Gly Ser Cys Cys His Gln Cys Lys Leu Leu 90 Ala Pro Gly Thr Leu Cys Arg Glu Gln Ala Arg Gln Cys Asp Leu Pro 105 110 Glu Phe Cys Thr Gly Lys Ser Pro His Cys Pro Thr Asn Phe Tyr Gln Met Asp Gly Thr Pro Cys Glu Gly Gly Gln Ala Tyr Cys Tyr Asn Gly Met Cys Leu Thr Tyr Gln 150 145

#### 25 Claims

5

10

15

20

50

- 1. A soluble meltrin polypeptide which does not comprise a transmembrane domain or an intracellular domain and which comprises the amino acid sequence of Glu (No. 156) to Ile (No. 686) from the N-terminal in Fig. 15a Fig. 15f.
- A polypeptide of claim 1 which comprises the amino acid sequence of Gly (No. 1) to Ile (No. 686) from the N-terminal in Fig. 15a Fig. 15f.
  - 3. A DNA comprising a base sequence encoding a polypeptide of claim 1 or 2.
- 4. A DNA of claim 3 which comprises the base sequence of No. 1 to No. 2058 from the 5' terminal in Fig. 15a Fig. 15f.
  - 5. A DNA of claim 3 which comprises the base sequence of No. 1 to No. 2848 from the 5' terminal in Fig. 15a Fig. 15f.
- An antisense oligonucleotide which hybridizes with a part of the sequence of No. 1957 to No. 2848 from the 5' terminal in Fig. 15a - Fig. 15f.
  - 7. An antisense oligonucleotide of claim 6 which inhibits the expression of the polypeptide of claim 1 or 2.
- 8. An antibody which recognizes the C-terminal region of a soluble meltrin wherein the C-terminal region is from amino acid No. 653 to No. 686 from the amino terminal in Fig. 15a Fig. 15f.
  - 9. An antibody of claim 8 which is a polyclonal antibody obtained from a mouse.
  - 10. 4. An antibody of claim 8 which is a monoclonal antibody.
  - 11. An antibody of claim 10 which is a monoclonal antibody obtained from a hybridoma using mouse spleen cells and lymphocytes.
  - 12. A method for the preparation of an antibody which method comprises:
    - immunizing an animal with a polypeptide of claim 1 or 2; and
    - obtaining an antibody from the immunized animal.

- 13. A vector comprising a DNA of any one of claims 3 to 5.
- 14. A transformant by the vector of claim 13.
- 5 15. A process for producing a polypeptide of claim 1 or 2, which process comprises culturing the appropriate transformant of claim 14.
  - 16. A medical composition comprising a polypeptide of claim 1 or 2, an antisense oligonucleotide of claim 6 or 7 or an antibody of any one of claims 8 to 11.
  - 17. Use of a polypeptide of claim 1 or 2, an antisense oligonucleotide of claim 6 or 7 or an antibody of any one of claims 8 to 11 for the manufacture of a medicament for treatment of a condition associated with unhealthy enhanced bone resorption.
- 15 18. Use according to claim 17, wherein the condition associated with unhealthy enhanced bone resorption is osteoporosis or hypercalsemia.
  - 19. Use of a polypeptide of claim 1 or 2, an antisense oligonucleotide of claim 6 or 7 or an antibody of any one of claims 8 to 11 for the manufacture of a medicament for preventing metastasis of cancer cells.

Patentansprüche

10

20

35

- Lösliches Meltrin-Polypeptid, nicht umfassend eine Transmembrandomäne oder eine intrazelluläre Domäne und umfassend die Aminosäuresequenz von Glu (Nr. 156) bis Ile (Nr. 686) von dem N-Terminus in Fig. 15a - Fig. 15f.
  - 2. Polypeptid nach Anspruch 1, umfassend die Aminosäuresequenz von Gly (Nr. 1) bis Ile (Nr. 686) von dem N-Terminus in Fig. 15a - Fig. 15f.
- 30 3. DNA, umfassend eine Basensequenz, welche ein Polypeptid nach Anspruch 1 oder 2 kodiert.
  - 4. DNA nach Anspruch 3, umfassend die Basensequenz von Nr. 1 bis Nr. 2058 von dem 5'-Terminus in Fig. 15a Fig. 15f.
  - 5. DNA nach Anspruch 3, umfassend die Basensequenz von Nr. 1 bis Nr. 2848 von dem 5'-Terminus in Fig. 15a Fig. 15f.
  - **6.** Antisense-Oligonucleotid, welches mit einem Teil der Sequenz von Nr. 1957 bis Nr. 2848 von dem 5'-Terminus in Fig. 15a Fig. 15f hybridisiert.
- Antisense-Oligonucleotid nach Anspruch 6, welches die Expression des Polypeptids nach Anspruch 1 oder 2 inhibiert.
  - 8. Antikörper, welcher die C-terminale Region eines löslichen Meltrins erkennt, wobei sich die C-terminale Region von Aminosäure Nr. 653 bis Nr. 686 von dem Aminoterminus in Fig. 15a Fig. 15f erstreckt.
- 45 9. Antikörper nach Anspruch 8, welcher ein polyklonaler Antikörper ist, der von einer Maus erhalten wurde.
  - 10. Antikörper nach Anspruch 8, welcher ein monoklonaler Antikörper ist.
- 11. Antikörper nach Anspruch 10, welcher ein monoklonaler Antikörper ist, erhalten aus einem Hybridom unter Verwendung von Mäusemilzzellen und Lymphozyten.
  - 12. Verfahren zur Herstellung eines Antikörpers, wobei das Verfahren umfasst:
    - Immunisieren eines Tieres mit einem Polypeptid nach Anspruch 1 oder 2; und
    - Erhalten eines Antikörpers aus dem immunisierten Tier.
  - 13. Vektor, umfassend eine DNA nach einem der Ansprüche 3 bis 5.

- 14. Transformant durch den Vektor nach Anspruch 13.
- 15. Verfahren zur Herstellung eines Polypeptids nach Anspruch 1 oder 2, wobei das Verfahren Kultivieren des geeigneten Transformanten nach Anspruch 14 umfasst.
- 16. Arzneimittel, umfassend ein Polypeptid nach Anspruch 1 oder 2, ein Antisense-Oligonucleotid nach Anspruch 6 oder 7 oder einen Antikörper nach einem der Ansprüche 8 bis 11.
- 17. Verwendung eines Polypeptids nach Anspruch 1 oder 2, eines Antisense-Oligonucleotids nach Anspruch 6 oder 7 oder eines Antikörpers nach einem der Ansprüche 8 bis 11 zur Herstellung eines Medikaments zur Behandlung eines Zustandes, der mit krankhafter gesteigerter Knochenresorption in Zusammenhang steht.
  - 18. Verwendung gemäß Anspruch 17, wobei der Zustand, der mit krankhafter gesteigerter Knochenresorption in Zusammenhang steht, Osteoporose oder Hyperkalzämie ist.
  - 19. Verwendung eines Polypeptids nach Anspruch 1 oder 2, eines Antisense-Oligonucleotids nach Anspruch 6 oder 7 oder eines Antikörpers nach einem der Ansprüche 8 bis 11 zur Herstellung eines Medikaments zur Verhinderung einer Metastasierung von Krebszellen.

#### Revendications

- Polypeptide de meltrine soluble, qui ne comporte pas un domaine transmembranaire ou un domaine intracellulaire et qui comporte la séquence d'acides aminés allant du résidu Glu n° 156 au résidu Ile n° 686 du domaine N-terminal présenté sur les figures 15a à 15f.
- Polypeptide conforme à la revendication 1, qui comporte la séquence d'acides aminés allant du résidu Gly n° 1 au résidu lle n° 686 du domaine N-terminal présenté sur les figures 15a à 15 f.
- 30 3. ADN comportant une séquence de bases qui code un polypeptide conforme à la revendication 1 ou 2.
  - 4. ADN conforme à la revendication 3, qui comporte la séquence de bases allant de la base n° 1 à la base n° 2058 du domaine 5'-terminal présenté sur les figures 15a à 15f.
- 35 ADN conforme à la revendication 3, qui comporte la séquence de bases allant de la base n° 1 à la base n° 2848 du domaine 5'-terminal présenté sur les figures 15a à 15f.
  - 6. Oligonucléotide anti-sens qui s'hybride avec une partie de la séquence allant de la base n° 1957 à la base n° 2848 du domaine 5'-terminal présenté sur les figures 15a à 15f.
  - Oligonucléotide anti-sens conforme à la revendication 6, qui inhibe l'expression d'un polypeptide conforme à la revendication 1 ou 2.
- 8. Anticorps reconnaissant un domaine C-terminal d'une meltrine soluble, lequel domaine C-terminal va du résidu d'acide aminé n° 653 3 au résidu n° 686 du domaine amino-terminal présenté sur les figures 15a à 15f.
  - 9. Anticorps conforme à la revendication 8, qui est un anticorps polyclonal obtenu chez une souris.
  - 10. Anticorps conforme à la revendication 8, qui est un anticorps monoclonal.
  - 11. Anticorps conforme à la revendication 10, qui est un anticorps monoclonal obtenu à partir d'un hybridome formé avec des cellules spléniques de souris et des lymphocytes.
  - 12. Procédé de préparation d'un anticorps, lequel procédé comporte :
    - le fait d'immuniser un animal avec un polypeptide conforme à la revendication 1 ou 2,
    - et le fait de récupérer un anticorps chez cet animal immunisé.

25

40

15

5

55

- 13. Vecteur comprenant un ADN conforme à l'une des revendications 3 à 5.
- 14. Organisme transformé avec un vecteur conforme à la revendication 13.

10

20

25

*3*0

35

40

45

50

55

- 5 15. Procédé de production d'un polypeptide conforme à la revendication 1 ou 2, lequel procédé comporte le fait de cultiver un organisme transformé approprié, conforme à la revendication 14.
  - 16. Composition médicale comprenant un polypeptide conforme à la revendication 1 ou 2, un oligonucléotide anti-sens conforme à la revendication 6 ou 7, ou un anticorps conforme à l'une des revendications 8 à 11.
  - 17. Emploi d'un polypeptide conforme à la revendication 1 ou 2, d'un oligonucléotide anti-sens conforme à la revendication 6 ou 7, ou d'un anticorps conforme à l'une des revendications 8 à 11, en vue de la fabrication d'un médicament conçu pour le traitement d'un état associé à une augmentation pathologique de la résorption osseuse.
- 15 18. Emploi conforme à la revendication 17, pour lequel l'état associé à une augmentation pathologique de la résorption osseuse est une ostéoporose ou une hypercalcémie.
  - 19. Emploi d'un polypeptide conforme à la revendication 1 ou 2, d'un oligonucléotide anti-sens conforme à la revendication 6 ou 7, ou d'un anticorps conforme à l'une des revendications 8 à 11, en vue de la fabrication d'un médicament conçu pour prévenir la métastase de cellules cancéreuses.

762 707 183	962 792	903
RGROUGISTA MEHGEMBIT BYDALFEEG BYDSGEIRGE DYGGLTVOIL BIILCLIENG FÜVYÜKAKTE MA <u>ni</u> stikkt twaklacine srifscerlo hh-bugsika <u>pehgingeri ba</u> ng labisdegae stelpvinv bivilvarn iragivitak aprojorsy aprejs henereleiten henggoche bandgeeg bi <u>dsge</u> ppps stetappist getkassen aligelilvi lighvicai cloipaeae bepazeage 1	oahhtrgkgi lærarhtyt pærhslucg mætsfilda ravpolgspo rvilplhyte hatt fassora <mark>dt f</mark> aspavroag girafsfice <mark>etf</mark> adfiskt glskplp ttrossl-pa krappdeset vstagfird aktrapppag pgavsss- <mark>21</mark> fapvxapkip agfræge <mark>ste 215</mark> ellærov	
L MAMETHER K APROTOBAS I LAMICATCA	BASPAVRQA BVPVXAPKI	
Frontanti Ingereran Alegines	PGAVESS-PI	
L BILCLIBA BLVILVABA OTTRASSENI	RVLLPLHGT	
Drigglavgii Stelpvswy Stetappert	ravpolospo Anprrp	<b>u</b> d
##55671R05 1.45V5020A	MODISKELDA VSTNORERPI	spreshrayi Valkve ior-
Pare Bress	Propression Trappopest D	APIRPERING GATOGEGER
METICENWIAE PETICHMODAS METICOGETES	LIGHAPHTHT TTRDSSL-PA PEEERAREE	PCOOEPOUR VIDETGETVE
RGPSVGFIDER NR-SVGFHKR NEWSTERISCO	Oakhtpakal alsnplf Leepepepe	splisalvrt pcoosponed apirogenho (spreshkayi k Kpterptep vægtegtve gatogeger (salkvpiok- r
72 Z	# # # B	# K

3

190 200 210 220 240 240 CGCGGGGGCCCCCCGAAGCAGCTGCCAGGCCGCGCGCGCG	250 260 270 280 290 300 conserve conser	×	360	TGCAGCCCGAGGGATGAGTTTGTGGGACCAGAGGGGGCTTACGAAGTGGCCAGAGCCTC	လ	420	CCTTCTGAGCAAGGACCCTGGGATCCCAGGACAGAGCATCCCAGCCAAGGATCATCCAGA	9	480	CGTGCTGACTGCAACTGCAGCTGGAGAGCCGAGACCTGATCCTCAGCCTGGAAAGGAA
66C A	ည	<u>a</u>		AGC	Y		TCC	Д		AAC
) P	360	¥		CAG	×		TCA	H		GGA
230 AGCG(	290 TGCT		350	၁၅၅	¥	410	GGA	2	470	S
E E	% 52 7.		က	IGT	>	4	JAA.	×	4	3AG
220 230 24 ACAATGGCAGAGCGCCCGGCGCC MAERPAR	Ş	A		GA	四		25	Ä		Ķ
ATG M	999	ري د		TAC	<b>&gt;</b> -	_	$\mathcal{Z}_{\mathcal{Z}}$	۵,	_	AT(
220 ACA	280 36CTG	Ä	340	GCT	A	400	ATC	I	460	STS
909	CTG			GGA			AGC	S		GAC
. වු	ည	A		AGA	84		CAG	O		<b>₹</b> 3
210 CAGG	270 GCTG	PARALLLALAGALLAPR	330	CAG	G M S L W D Q R G A Y E V A R A S	390	GGA	K D P G I P G Q S I P A K D H P D	450	AGC
200	2 CTG		က	GAC	<u>a</u>	က	CCA	<b>A</b>	4	GAG
CAC	STC			766	E=		ATC	<b>H</b>		CTG
CTG.	Ç.			ITG.		0	366,	(5)	0	CAG
200 .AGC	260 GCGC	~	320	GT	<b>-</b>	380	SCT	٥	440	CTG
AGC	S			TG/	· • ·		;ACC	_		SAA(
<b>₹</b> 50				795		•	AGG	,—		57.
190 CCCC	250 CCC	) 	10	GAG	9	370	SCA	×	430	CTC
¥ 2395	<b>72</b> 22	2	310	ဗ္ဗ	×	က	<b>IGA</b>	S	せ	IGA
35	יט	R A P	•	;AG	A A R		TC	S 7 7		5
252	S. S.	} ≈		TGC	4		S	7		5

FIG.2b

107	127	147	167	187
540 .TGT V	600 TGC A	660 LAAA N	720 AGC A	780 CAC T
D G	GA'	E GA	, )) P (C,	, Š
ract T	G	F	SGTC V	SAAC
530 ATGG1	590 TGCAA	650 TCAT	710 AACT(	770 AGTC(
S. AGA. D	59 FGTC V	69 FAT( I	CAA.	ZAA(
GCA/	ACA:	ACT.	CTA(	TAAC
O TCT	0 TGG, G	7555 0	O CAGO	GCA'
520 TTATC Y	580 CCATC H	640 CCGGG R (	70 TGA D	760 7CACAGC S Q H
CCA H	CTA	TCT	700 ACCACTGAC T T D	GTC
3AC T	ΓΤΑ Υ	Š O	CAC	ည်
490 500 510 520 530 540 TGAGGGACTCATTGCCAATGGCTTCACGGAGACCCATTATCTGCAAGATGGTACTGATGT E G L I A N G F T E T H Y L Q D G T D V	550 560 570 580 590 600 CTCTCTCACTCGAAATCACGGATCATTGTTACTACCATGGACATGTGCAAGGAGATGT S L T R N H T D H C Y Y H G H V Q G D A	610 620 630 640 650 660 TGCATCAGTGCTCAGTACTTGCTCTGATCTCCGGGGACTTATCATGTTTGAAAA A S V V S L S T C S D L R G L I M F E N	670 680 690 700 710 720 TAAAACGTACAGCTTAGAGCCAAAAAAAAAAAAAAAAAA	730 740 750 760 770 780 TGAGAGCATGACAAGGGCTGTGTGGGTCACAGCAAGTCCAACCTCAC E S M T N I Q G L C G S Q H N K S N L T
CAC →	TCA	TTG	GAA	125 1-
CIT	GGA	TAC	AAT	AGG G
500 ATGG	560 ACAC	620 TCAG	680 AGCC	740 TCCA
CAA	TCA H	CCT	AGA E	CAT I
TGC A	AAA N	CAG S	CIT	GAA
CAT	O TCG	0 GGT V	O CAG	GAC T
490 ACTC	550 CACT T	610 AGTGC	670 GTACA	730 CATGA
490 GAGGGACTCATTGCC E G L I A	TCT	ATC	AAC	GAG
TGA	CTC	TGC	TAA	TGA

	J
(	1
C	j
L	4

	207	227	247	267	287
840	JAC T	860 870 880 890 900 AAGTACGTAGAGTTATTGTGGCAGACACAGAGAGTTTCAGAG K Y V E L V I V A D N R E F Q R	960 TGA D	1020 ATGA I D	1080 ACTG
<b>~</b>	E GA	Š Š o	SGT)	3AA.	AGA(
	TCCCCTGGAACCTCTCAAATGCGGCAAGAAGGCATAAGAGAGAG	FF	910 920 930 940 950 96( CAAGGAAAAGACTGAAGCAGCGATTAATAGAGATCGCCAATCACGTTG, Q G K D L E K V K Q R L I E I A N H V D	980 990 1000 1010 1020 CCACTGAACATCCGGATCGTGGTGGGAATGA P L N I R I V L V G V E V W N D	1040 1050 1060 1070 108 TCTATAAGCCAGGACCCATTCACCAGGCTCCATGAGTTTCTAGACT S I S Q D P F T R L H E F L D W
830	ľAA( K	890 GAGA(	950 CCAA	10 AGT( V	1070 GAGTTI E F
œ	E E	CAG.	6 V	1010 GGAAG1 E V	10 TGA
	AAG R	N CAA	GAT(	ÄĞT	CCA
0	AAG. R	AGA(	O AGA( E	AGG.	o GCT L
820	36C/ A	880 36CA(	940 AATA( I	1000 GGTAG	1060 CAGGC R L
	<u>පි</u> දි	rgt(	ATT. L	GCT	CAC
	AAT( M	ľAľ I	GCG	CGT V	ATT
810	چ م	870 GGT	930 GCA(	990 GAT( I	1050 GACCCA D P
	STC	ic I	FAA	ر ه	1 GGA D
	T T	\GA(	\GT	CAT	. SC &
008	7991 G	860 ACGT/	920 AGAA/	980 TGAA(	1040 ATAAGCC I S Q
<b>∞</b>	)    -	86 GTA(	GGA(	ACT	10, TAT, I
	S S	CAAC	CC I	ACC, P	CTC
0	rg. V	3AC	O AGA D	CAG. R	O ATG
790	ATGGAAGATGI M E D V	850 CTTAAGATGAC L K M T	910 \AAAG K L	970 AAGTTTTACA( K F Y R	1030 ATCGACAAATC I D K C
	Šč ≅	raac K	.99\ .0	FF	D GA
	CATGGAAGATGTCTCCCTGGAACCTCTCAAATGCGGCAAGAAGGCATAAGAGAGAG	850 860 870 880 890 900 CCTTAAGATGACCAAGTACGTAGGGTTATTGTGGCAGACAACAGAGAGTTTCAGAG	910 920 930 940 950 960 GCAAGGAAAAGACAGTTAAGCAGTTAATAGAGATCGCCAATCACGTTGA Q G K D L E K V K Q R L I E I A N H V D	970 980 1000 1010 1020 CAAGTTTTACAGACCATCGATCGTGCTGCTGGAGTGTGGAATGA K F Y R P L N I R I V L V G V E V W N D	1030 1040 1050 1060 1070 1080 CATCGACAAATGCTCTATAAGCCAGGACCCATTCACCAGGCTCCATGAGTTTCTAGACTG I D K C S I S Q D P F T R L H E F L D W

-	The state of the
(	1
(	5
۲	_
Ţ	Ι.,

307	327	347	367	387
1140 GGGT	1200 AACA	1260 TGGC	1320 GCTG C	1380 TCCC
11 366 3	12 GAA E	12 TTG L	13 AGC S	13 TTC F
1100 1110 1120 1130 1140 TCTACCTCGAAAATCCCACGACAATGCTCAGCTTATCAGTGGGG L P R K S H D N A Q L I S G V	1160 1170 1180 1190 1200 CCACCATGGCACCCATCATGAGCATGTGCACTGCAGAAC T I G M A P I M S M C T A E Q	1220 1230 1240 1250 126 ATGGACCATTCAGACAGCCCCTTGGTGCCGCAGTGACCTTGG M D H S D S P L G A A V T L A	1270 1280 1290 1300 1310 1320 CATGAGCTGGGCCACAACTTCGGGATGAACCATGACACCACTGGAGGGGCTGCAGCTT H E L G H N F G M N H D T L E R G C S C	1340 1350 1360 1370 1380 SAGAAAGGAGGCTGCATCATGAACCCGTCCACGGGGTTCCCATTCCC E K G G C I M N P S T G F P F P
VTC.	VCT.	TG.	395 3	o ITC F
1130 CTTA1 L I	.1190 .crccac .c T	1250 GCAG1 A V	1310 3AGGGG R G	.1370 GGGGT1 G F
AGC L	TGT	CC 97	AGA	7 9 0
TC.	3CA	GTG A	TGG E	CCA
:0 VTG( A	30 rgA(	1240 CCTTGG1 L G	1300 CACAC T L	1360 CCCGT P S
1120 CAATO N	1180 CATG/	12, CCJ L	13( VCA(	13. P
CGA	CAT	ည် ရ	TG/ O	ZGA/
CCA	ACC P	CAG S	CCA	CA1
1110 AATC(	1170 TGGCA I A	1230 CAGA(	1290 TGAA(	1350 GCAT(
AA X	11 XAT(	11. TC/	15 N M	ر ت يق ت
CGA R	9	CA7	999	9
CCT P	O ATC İ	O GAC D	O TTC F	0 0
1100 CTAC L P	1160 ACCA T I	1220 ATGG/ M D	1280 AACT N F	1340 AAAG6 K G
TTC	CC		AC.	AG.
AGC	GAA	TTG	00 H	CAG
AA K	50 AAG G	10 3AG V	70 766 9	1330 GGCCG A A
1090 GATA/ I	1150 CCAA(	1210 GAGGAC G V	1270 \GCTG( L (	133 PGG A
1090 AGAAAGATAAAGC R K I K L	1150 TATTTCCAAGGAA Y F Q G T	1210 TCTGGAGGAGTTG S G G V V	16/ E	1330 AGAATGGCCGCA R M A A
1090 1100 1110 1120 1130 1140 GAGAAAGATAAAGCTTCTACCTCGAAAATCCCACGACAATGCTCAGCTTATCAGTGGGGT R K I K L L P R K S H D N A Q L I S G V	1150 1160 1170 1180 1190 1200 TTATTTCCAAGGAACCACCATCGCCATCATGAGCATGTGCACTGCAGAACA Y F Q G T T I G M A P I M S M C T A E Q	1210 1220 1230 1240 1250 1260 GTCTGGAGGAGTTGTCATTGGACAGCCCCCTTGGTGCCGCAGTGACCTTGGC S G G V V M D H S D S P L G A A V T L A	1270 1280 1290 1300 1310 1320 ACATGAGCTGGGCCACTTCGGGATGAACCATGACACTGGAGAGGGGCTGCAGCTG H E L G H N F G M N H D T L E R G C S C	1330 1340 1350 1360 1370 1380 CAGAATGCCCCAGAGAAAGGAGGCTGCATCATGAACCCGTCCACGGGGTTCCCATTCCC R M A A E K G G C I M N P S T G F P F P
9	<b>.</b>		•	•

FIG.2e

	407	427	447	467	487
1440	UNIGUIGITICAGCAGCAGGAAGGACCTGGAGGGCATGGG	1500 GAAA N	1560 ATCG	1620 AGTG	1680 CCTG C
<del>-</del>	CAT	1. 166, 6	N N	CAC	16 TCC S
	တ္တ် ၁	ည်း	ACG T	9 999	AAC
30	SAA( K	30 SAAG K	0 757 0	CAC H	O AGC S
1430	3GA E	1490 CCGCA/ R K	1550 GAAT( E C	1610 GCGC/ A H	1670 TCCAG S S
	्रा	ည်း	GAG E	75C C	ည်
	S	999	0 0 1	GTG V	AGG R
1420	ည် န	1480 CTTT F (	1540 AGAA( E	1600 TGCTG A V	1660 ATGC/ C
	S E	A A	1 GGA G	1 GAT D	GCA'
	) 1	o O	736C C	CCA P	ACT
1410	දිද <del>ි</del> ර	1470 TCAA(	1530 GTGA(	1590 TGAAGC K P	1650 CAGGA G
7	SA K	GGT V	11. 3.TG	ر ت ت ت	CCA P
. (	AGCAGCTGCAGGACCTGGAGCCTGGAGAGGGCATGG S S C S R K D L E A S L E K G M G	AGA(	AGA(	1580 1590 1600 1610 1620 ACCACCTGTACTCTGAAGCCAGATGCTGTGTGCGCCGCCACGGGCAGT T T C T L K P D A V C A H G Q C	1640 1650 1660 1670 1680 GTCAGCTGAAGCTGCATGCAGGGGCTCCAGCAACTCCTG
1400	SAG	50 ACC	20 NGAA E	1580 ACCTGT T C	EO FAAG K
14	ည် ည	1460 CCTACC L P	1520 366AGA G E	158 2ACC T	1640 CTGAA L K
,	S	CAA	AGA(	racc T	CAG 0
0	S	1450 1460 1470 1480 1490 150 ATGTGCCTCTTCAACCTACCAGGGTCAAGCAGGCCCTTTGGGGGCCCGGAAGTGTGGAAA M C L F N L P E V K Q A F G G R K C G N	GGA. E	SGC; A	ر د ترون
1390	7 F	1450 CCTCT	1510 TGTGG V E	1570 TGCTGTAACG C C N A	1630 TGTGAAGACT C E D C
Ę	Algeleri M V F	GTG C	ZTA'	CYTG]	Ç. Ā.
	3	1450 1460 1470 1480 1490 1500 GATGTGCCTCTTCAACCTACCAGGTCAAGCAGGTGTGGAAA M C L F N L P E V K Q A F G G R K C G N	1510 1520 1530 1540 1550 1560 TGCTATGTGGAAGAGGAGAGGTGTGACTGCGGAGGAATGCACTGG G Y V E E G E E C D C G E P E E C T N R	1570 1580 1590 1600 1610 1620 CTGCTGTAACGCTACCACTGTACTCTGAAGCCAGAGTG C C N A T T C T L K P D A V C A H G Q C	1630 CTGTGAAGACT C E D C

FIG.2f

507	527	547	292	587
1740 TACA	1750 1760 1770 1780 1790 1800 TGATGGCCACCCGTGTGGTTACTGCTACAACGGCATCTGCCAGACCCA D G H P C Q G V D G Y C Y N G I C Q T H	1860 GCTT , F	1920 GCGC	1980 GTGC
1690 1700 1710 1720 1730 1740 TGACCTCCCAGGAATTCTGCACCAGGGACTGCCCCTCACTGTCCCAATGTGTACCTACA D L P E F C T G T A P H C P A N V Y L H	1750 1760 1770 1780 1790 1800 GATGGCCACCCGTGTGGTTACTGCTACAACGGCATCTGCCAGACCC D G H P C Q G V D G Y C Y N G I C Q T H	1810 1820 1830 1840 1850 1860 TGAGCAGCAGTGTCACGCTCTGGCATCTGCTT E Q Q C V T L W G P G A K P A P G I C F	1870 1880 1890 1900 1910 1920 TGAGCGAGTCAACTGTATGGTAACTGTGCCAAAGACTCCAAGAGCGC E R V N S A G D P Y G N C G K D S K S A	1930 1940 1950 1960 1970 1980 CTTCGCCAAATGTGAGAGATGCCAAGTGTGGGAAAATCCAGTGTCAAGGTGGTGC F A K C E L R D A K C G K I Q C Q G G A
GTA( Y	رر <del>V</del>	CAT	CAA	AGG
1730 AATGT( N V	1790 ATCTG I C	1850 CCTGG P G	1910 GACTC D S	1970 TGTCAA C 0
17 CAA N	17 CAT I	18 17CC	19 IAGA D	19 NGT C
AGC A	ອ	995 A	SCA.	722
TCC P	N CAA	FO P	1900 CTGTG( C G	1960 GAAAA1 K I
1720 ACTGT( C 1	1780 3CTAC/	1840 CTAAA( K	19( ACT(	19( GGA/ K
CTC/ H	ACT	GŤG( A	GTA. N	GTG G
CC A	STT	O CAG	O A TG	O. AGT
1710 ICTGC(	1770 ATGGT	1830 GACC/	1890 CTTA	1950 CCAAG K
GGA(	166/ D	5	ATC P	ATG
CAG G	) ) )	TCT	GAG D	GAG
1700 TGCA( C T	1760 CAGGG	1820 ACGC T L	1880 GCAGG A G	1940 CTGA( L R
1 TCT C	1 GTC	1 TCA	CTG A	1 1 1 1 1 1 1 1 1
AAT F	CGT	GTG V	ACT	CTC
1690 GACCTCCCAGAAT D L P E F	1750 CCACC H P	1810 GCAGT 0 C	1870 AGTCA V N	1930 TTCGCCAAATG F A K C
16 TCC P	17 GCC H	18 AGC	18 GAG	19 10CA
ACC T	ATG G	AGC	AGC	
TG	TG	T E	TG B	2.

209	627	. 647	199	687
2040 AGGA	2100 ACCC	2160 3GATG	2220 iGGGT	2280 ACAA K
20 CAG	21 GAC D	% CG 21	. G 22	GA D
CAG	CCA P	۳. ۳	~ 790 ₩	CTG C
2000 2010 2020 2030 2040 FCATTGGTACCAATGCTGTTTCCATAGAAAATATCCCACAGGA I G T N A V S I E T N I P Q Q E	2060 2070 2080 2090 2100 CTGTGCCGGGGGACCCATGTTGGTTGACATGCCAGACCO	N N	0 0 0 0	2230 2240 2250 2260 2270 2280 ATGTAACAACAGAAGAATTGCCACTGTGAAGCCCACTGGGCTCCACCCTTCTGTGACAA CON N R K N C H C E A H W A P P F C D K
2030 ATCCC I P	2090 IGACA1 D M	2150 CCTCAA L N	2210 CCACGG H G	2270 ACCCTT P F
N	ĜA1 D	orte C	3TG C	TCC
T		AAT(	GCA(	0 GGC
2020 'AGAA/ E	2080 CTTG( L (	2140 !AAAAA K I	2200 CATGC M Q	2260 CTGGG
ZATA I	STA(	AGG/	TGC A	⊞ CCA
rtc( s	·	AGA. E	GTG	AGC A
)10  GT7	2070 CCCA7	. 2130 FTGTGC/ C A	2190 :ACAAG I K	2250 GTGA.
2010 TGCTGT	24 3AC(	616° C	rca H	2 CTG
SAA7	9	AAA(	CGT	CCA
TACC	30 R R	2120 GGAAC, G T	ဝ <u>္</u> ထ ၁၁	2240 !AATTG .N C
2000 rggta( g t	2060 3TGCCG C R	213 AGG, G	2180 CTTCG( F G	22 GAA .N
CAT	- 15 1	TGC.	TGT V	GAA
rgt(	O GAT	o CCT	O CAG	CAG R
JG. 28 1990 AGCCGACCTGT S R P V	2050 GGAGGTCGGAT G G R I	2110 FTGTGC V L	2170 NTATCA I S	2230. 1CAACA N R
37 . SCG.	AGG.	GCT	GAA	TAA
FIG. 28 1990 AAGCCGACCTGT S R P V	2050 2060 2070 2080 2090 2100 AGGAGGTCGGATTCTGTGCGGGGGGGGGCCCATGTGTTTGGGTGATGACATGCCAGACCC G G R I L C R G T H V Y L G D D M P D P	2110 2120 2130 2140 2150 2160 AGGGCTTGTGCTGCAGGAACAAAATCTGCCTCAATCGTCGATG G L V L A G T K C A E G K I C L N R R C	2170 2180 2190 2200 2210 2220 TCAGAATATCAGTGTTCACAAGTGTGCCATGCCAGGGCGGT Q N I S V F G V H K C A M Q C H G R G V	ATG C

4	Į
(	
C	j
Ц	_

707	727	747	767	787
2290 2300 2310 2320 2330 2340 GTTTGGCTTTGGAGGAAGCACAGAGCAGGATAACCAGGGCTT F G F G G S T D S G P I R Q A D N Q G L	2360 2370 2380 2390 2400 CCTGGTGAGCATCCTGCTTGCTGCTGCTGTGTTTTGTGGTGTATCT L V S I L C L L A A G F V V Y L	2460 AAAGCT K L	2520 CCACAC H T	2580 CAGGCA R H
Č Č	GGT	GGA	TCA( H	GGA(
2300 2310 2320 2330 234( GGAAGCACAGACGGTCCCATCAGGCAGCAGATAACCAGGGCT1 GSTDSGPIRQADNQGL	2350 2360 2370 2380 2390 2400 GACTGTAGGAATCCTGGTGAGCATCCTGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGTGTATCT T V G I L V S I L C L L A A G F V V Y L	2410 2420 2430 2440 2450 2460 CAAAAGGAAGACGTTGATGCTGCTGTTCACATAAAAAAAA	2470 2480 2490 2500 2510 2520 AAGGTGTGTGCCCCTTCCCGGACACCCAGGCCCACCACCACCACCACCACCACCAC	2530 2540 2550 2560 2570 2580 CCCCGGGAAAGGCCTGATGAACCGGGCACCACATTCCATACCCCCAAGGACAGGCA P G K G L L M N R A P H F N T P K D R H
A A	TGG G	T	1 1 1	TAC
2320 CAGGC/ R Q	2380 TGCTGC A A	2440 TAAAAA K K	2500 TCACCT H L	2560 TTTCAA F N
CCATC	2 rectt L	2 CACAT H	3cccT	2 CACAT H
O GTC P	0 GTC L	0 TCA(	o GTG( G	O CAC
2310 CAGTGG; S G	2370 CCTGTG1 L C	2430 SCTGTT( L F	2490 ACCCAG P S	2550 CGGGC/ R A
AGA(	CAT( I	GCT( L	GAC	GAAC
2300 AGCAC S T	2360 GTGAG V S	2420 ATGCG M R	2480 TCCCGGACA S R T	2540 CTGAT
23 3AAG S	23 rGGT V	24 IGAT	24 7TTC S	25 16CT
G. G.	[] []	LL 50 LL 50	222	<u> </u>
2290 CTTTG( F G	50 GAA'	10 AGA(	70 TGC/ H	2530 GAAAG( K G
22 GCT F	2350 TAGGA/ G ]	2410 GGAAGA K T	2470 GTGTGC V H	25; 36A/ K
2290 TTTGGCTTTG F G F G	2350 ACTGTAGGAA' T V G I	2410 AAAAGGAAGA( K R K T	AGGTGTGTGC/ R C V H	2530 CCCGGGAAAGG P G K G
F.	GA	CA	AA R	S

282	867	847	827	807
2880 GACC	2820 CTAG S	2760 AACC	2700 GTGG	2640 CACA
28 AGA	28 ACT T	27 AAA K	27. XAG7 S	200 A
ATC	CTC	CGA	ر الاردر 1	SGT(
	0 درن	0 ATT I	O A A A	A A
2870 GCCCC	2810 TCTCG S R	2750 GGCA1 G I	2690 .cgtg( R A	2630 'CGAG( R A
CCA	ACT	CAG	P P	SGCJ A
	AGG,	000 V	ACC	GAC D
2860 GCAC(	2800 GAGCA S F	2740 GCAGG	2680 CCAG/	2620 CCTC( L
366	CTG	AGG R	CAC H	CCC P
. CT	CCA(	GTC V	CTC	AGG R
50 3AA(	2780 2790 2800 2810 2823 3AAGCCTCTGCTGATCCACTGAGCAGGACTTCTCGGCTCACTA K P L P A D P L S R T S R L T S	2720 2730 2740 2750 2760 CCTGCCGCCAGTCCAGGCAGGCCCAGGGCATTCGAAAAC L P A S P A V R Q A Q G I R K P	2660 2670 2680 2690 2700 CAGCGAGTGCTCCTCCCACCAGACCCCAGTGCACCCAGTGC Q R V L L P L H Q T P R A P S G	2600 2610 2620 2630 264 CAGAACATGGACATCAGCCGCCCTCGAGCCGTCCCAC Q N M D I S R P L D A R A V P Q
2850 CAGGA/	27. GCT	27 CCT	26 CTG L	26 ATC 1
CAG	CCT	AGT	CTC	GAC D
9999	2780 CCTCTG( P L I	0 6 6 6	0 GTG V	O ATG M
2840 CCAGG P G	278 CCT	2720 CCCGC P A	2660 CGAG7 R V	2600 AACA1 N M
NCC H	YAG	CTG	CAG Q	CAG
7951	CAG/	200	r P	_ C C
2830 'GGTGA V F	2770 TCCTC	2710 CAGGC R P	2650 GTCAC S P	2590 GAAA1 K C
7.00 V	27CTC	2. 3CC/	CA S	2 CTG L
2830 2840 2850 2860 2870 2880 TGCCTTGGTGAGGACCCCAGGCAGCCAGCCCCATCAGACCCATCAGACCCAGGCAGG	2770 2780 2790 2800 2810 2820 CAGTCCTCCTCAGAGCCTCCTGCTGAGCAGGACTTCTCGGCTCACTAG S P P Q K P L P A D P L S R T S R L T S	2710 2720 2730 2740 2750 2760 CCCTGCCAGGCCCCTGCCAGTCCAGGCCAGGCCCAGGGCATTCGAAAACC P A R P L P A S P A V R Q A Q G I R K P	2650 2660 2670 2680 2690 2700 GCTTCAGTCACCTCAGCCTCTCCAGCCCAGCGTGGG L Q S P Q R V L L P L H Q T P R A P S G	2590 2600 2610 2620 2630 2640 CTCGCTGAAATGCCAGAACATGGACATCAGCCGCCCCTCGACGCTCGAGCCGTCCCACA S L K C Q N M D I S R P L D A R A V P Q
TG	S	<u>ي</u> د	· 35 T	. U"

FIG. 2j

20	40		80	100
120 TAC	180 GCC P	240 IGAG	300 TAG	360 CAA
SCAGC Q L	AGGTA G S	rggga g R	SATCC	CACTG
110 GCTCT( A L	170 CAAGAA Q E	230 LAGCCC S P	290 GAGCT( E L	350 TGCTA C Y
80 90 100 110 120 XGCGGGCGTCCCCGGTTCTGCTTGCTCGCTCTGCTAC A G V A R F C L L A L A L Q L H	130 140 150 160 170 180 TGGCCGCTGGCGGGGGGGGGGCCAAGAGGTAGCC W P L A A C E P G W T T R G S Q E G S P	200 210 220 230 240 GAACTCATAATACCTCAGTGGCGGACTTCAGAAAGCCCTGGGAGAG E L I I P Q W R T S E S P G R G	260 270 280 290 300 AGAGCAGAGCTCAGGCTGAAGGGCCGAGAGCTGATCCTAG R A E L R V M A E G R E L I L D	320 330 340 350 360 GAGCACCTTTTTGCTCCAGCCTACACAAACCTGCTACACTGCAA E H L F A P A Y T E T C Y T A S
) 166C1 A	3 3AGG/ G	o CTTC/ S	O AAGG( G	O CAGA/ E
100 TTGCTG L L	160 ACAAG/ T R	220 CGGACT1 R T S	280 GCTGAA A E	340 TACAC
CTGC	GACC	GTGG ₩	CATG	AGCC A
90 366TT	150 3GATG	210 ICCTCA P Q	270 16667 R V	330 SCTCC
CGCCC	) P (	AATA( I I	GCTC/ L 1	TTTT F
80 GCGT V V	140 TGCGA C E	200 CTCAT	260 GCÁGA A E	320 CACCT H L
) SCGG	1 16667 A	2 GAAC E L	2 ZAGAG R A	SGAGC E F
36CGC	reece A	NGCA1	CACTO	O AGAA( N
7. 3a 70 <u>ATG</u> CCCGGGC M P G R	130 CCGCTG P L	190 CCGCTACAGC, P L Q H	250 AAGCATCCAC K H P L	310 CTGGAGAAGA L E K N
TO 80 90 100 110 120  TCATGCCCGGGCGCGCGCCCCGGTTCTGCTTGCTCGCTCTCGCTCTGCTAC  M P G R A G V A R F C L L A L A L Q L H	130 140 150 160 170 180 ATTGGCCGCGCGCGGGGGGGGAGCCAAGAAGGTAGCC W P L A A C E P G W T T R G S Q E G S P	190 200 210 220 240 CTCCGCTACACATCATAATACCTCAGTGGCGGACTTCAGAAAGCCCTGGGAGAG P L Q H E L I I P Q W R T S E S P G R G	250 260 370 280 290 300 GAAAGCATCCACTCAGGGCCAGAGCTCAGGGTCATGGCTGAAGGGCGAGAGCTGATCCTAG	310 320 330 340 350 360 ACCTGGAGAAGGACCACCTTTTTGCTCCAGCCTACACAGAAACCTGCTACACTGCAA L E K N E H L F A P A Y T E T C Y T A S
<b></b>				

2_	ב
7	)
C	)
H	

120	140			200
400 410 420	460 470 480	520 530 540	580 590 600	640 650 660
CTGAGGATCACTGCTTTTACCACGGGACTG	CCACCTGCCGGGGATTAGAGGACTGA	CGAGCCCGTCCCTAACAGCGACAC	GCCCCCGGGGAACTGTGGGTTCGAGCACT	AGTTTACACATCAGACCAAAAGCAACCTC
E D H C F Y H G T V	S T C R G I R G L I	E P V P N S D S Q H	P P G N C G F E H S	F T H Q T K K Q P R
380 390	440 450	500 510	560 570	620 630
CAGCACGCTGAAGT(	GTCCAGTGTCACGCT	CCTCAGCTACATCA1	CAACATCTCACGC	GACTGGGCCCTTCA
S T L K S	S S V T L	L S Y I I	EHLTL	DWALQ
370 GTGGCAATCCTCAAACO	430 TGAGGGACGTGGATGA( R D V D E	490 TTAFAGTGAGAAGTAAC I V R S N	550 ACCGTATTTACAGATCC R I Y R S	610 CCGGGCCCACCTCGAAG G P T S K

☎.

220	240		280	300
720 CTG	780 .TGG	840 iTCG	200 200 006	960 CTC
099 V	CA Z	TG.	TA(	TG.
GGT V	GCT	ACT L	CTC S	Z &
CTCT	CAA K	D FGC, A	O CTA Y	O TGA D
710 TAC( Y 1	770 NCGC/ R	830 VATT( I	890 7000 P	950 CCATG H L
icr L	CAAA.	20G/R	GAA7 N	GAG( S
GA(	T.	AT( I	IGA(	K K
700 CGTG V	CAGAAGAATCGACATGACCAGGATGCCACCAAACGCAAGCTCATGG Q K N R H D Q D A T K R K L M E	800 810 820 830 840 GTTGATAAGTTTTACCGCTCCCTGAACATCCGAATTGCACTTGTCG V D K F Y R S L N I R I A L V G	860 870 880 890 900 ACGCATGGGGATAAGTGTGAAGTTTCAGAGAATCCCTACTCTACCC T H G D K C E V S E N P Y S T L	920 930 940 950 960 AGTTGGAGGGGCAAGCTGCTTGCTCAGAAGAGCCCATGACAATGCTC S W R R K L L A Q K S H D N A Q
TAC Y	GA7	, 5 , 1	GT.3	P P P
iÁAG K	CAG	S. S.	GAA E	L J
0 ATG M	GAC D	20 R R	0 1. 1. 1. 1. 1.	930 VAGCTG
690 TCTA' S M	750 CATG/ H D	810 TACCO	870 FAAGT( K C	93 XAG
ICAC H	icca R	FF	GA1	) ) ) (
CTA	SAA1	CAAC K		O GAĞC R
680 VGAT( D 1	740 GAAG/	800 FGAT	860 3CAT( H	920 ΓΤGG/
B. B.	ر م	· V	3AC(	rag'
ار م	FIT.	TAT Y	JTG(	ر 1
670 GAAA K	730 TATGCAGAGTT Y A E F	790 CAAC N	850 GGTG	910 TGGTCCTTTCT W S F L
ATC	GC/	A GC 7.	GAK E	STC
670 680 690 700 710 720 GCAGAATGAAGGTCTACACTCTATGAAGTACGTGGAGCTTTACCTGGTGGCTG R M K R E D L H S M K Y V E L Y L V A D	730 740 750 760 770 780 ATTATGCAGAGTTTCAGAAGAATCGACATGACCAGGATGCCAAACGCAAGCTCATGG Y A E F Q K N R H D Q D A T K R K L M E	790 810 820 840 AGATTGCCAACTAGTTGATAAGTTTACCGCTCCCTGAACATCCGAATTGCACTTGTCG I A N Y V D K F Y R S L N I R I A L V G	850 860 870 880 890 900 GCTTGGAGGTGTGGAGGATAAGTGTGAAGTTTCAGAGAATCCCTACTCTACCC L E V W T H G D K C E V S E N P Y S T L	910 920 930 940 950 960 FCTGGTCCTTTCTTAGTTGGAGGCGCAAGCTGCTTGCTCAGAAGAGCCATGACAATGCTC
3	A	AG	3	7.

_	Ċ	j
C	ř	)
(	-	)
	1	

970	6		980	00		6	066		Ä	1000			1010	0		Ξ	1020	
AGCTAATCACGGCAGGTCCTTCCAAGGCACCACTTGGCCTGGCCCCCTCATGGCCA L I T G R S F Q G T T I G L A P L M A M	<u>ن</u> وي	SCAC R	S. S	CTT F	GGGCAGGTCCTTCCAAGGCACCACTTGGCCTGGCCCCCTCATGGCCA G R S F Q G T T I G L A P L M A M	AGG G	CAC	CAC	CAT I	766. G	CZ L	ეე ▼	)) 	CCI L	CAT	<u>ک</u> و ۲	CCA	320
1030	<u> </u>		1040	9		. 01	1050		Ä	1060			1070	0		=	1080	
TGTGCTCCGTGTACCAGTCTGGAGGAGTTAGCATGGACCACTCCGAGAATGCCATTGGTG	rgt/	222	S S	.16G	GTACCAGTCTGGAGGTTAGCATGGACCACTCCGAGAATGCCATTGGTG Y Q S G G V S M D H S E N A I G V	AGT' V	TAG S	CAT	GGA(	CCA(	S	<b>β</b>	GAA	TGC	CAT I	प्रति क	3TG V	340
1090	<u> </u>		1100	Ó		. 1	1110		-	1120			1130	0		<del>-</del>	1140	
TAGCCTCCACTGTGGCCCATGAGATTGGCCACAACTTTGGCATGAGCCATGATTCTGCAC	TGI	) ) ) )	SC.	TGA	GAT	TGG	CCA:	CAA	CTT	16G	CAT	GAG	S	TGA	TTC	TG	CAC	
7 0 V	>	∢	E	z)	V A H E I G H N F G M V H U V A H	<u>ح</u>	I;	Z	I,	ح	Z	2	I	<b>a</b>	<b>S</b>	₹ .	I	360
1150			1160	0			20	1170 1180		180			1190	0		ä	1200	
<b>ACTGCTGTTCTGCCAGTGCCGATGGCGGCTGCATCATGGCCGCCGCCACCGGGCACC</b>	FIG.	)CA6	TGC	AGC.	CGA	TGG	999	CTG	CAT	CAT	395	3	33	CAC	993	35.	ACC	
ن د	A	S	A	A	A S A A D G G C I M A A A T G H P	ර	ِ ق	ပ	H.	×	A	¥	¥	H	೮	I	٠	380
1210	_		1220	0		12	1230		H	1240			1250	0		===	1260	
CTTTCCCCAAAGTGTTCAGTTGGTGTAACAGGAAGGAGGTGGACAGGTATCTGCAGACAG	<b>1AG</b> 3	[GT]	CAG	TTG	GTG	TAA	CAG	GAA	GGA	SCT	3GA	CAG	GTA	TCI	CCA	GAC	CAG	
F P X	>	ᄄ	S	æ	VFSWCNRKELDRYLQTG	Z	2	<b>~</b>	Œ	_	9	8	<b>X</b>		0		ပ	400

(	ט
7	$\mathbf{C}$
	•
(	7
$\vdash$	4
Ľ	Ļ

420	440	460	480	200
1320 AGGT R C	1380 TGTA C K	1440 CATG H G	0 D %	60 TT Y
13 GAG R	13 ATG C	CCA H	1500 SCAGGTTC Q V R	1560 CTATT Y Y
. Š	GGA, E	rgc A	3CA(	CAA
O AGG	O GGA(	GTG.	GGA(	CAC
1310 TGGAG G G	1370 AGAGGA( E E	1430 AGAGT E C	1490 TCGGGA R E	1550 CCCCA P T
3TA' Y	NGA.	A A	riei C	CTGC
3CT( L	1991	9	SCA(	H CAC
1300 GGAC(	1360 ACTG1	1420 AGGAA( E (	1480 GGAACC G T	1540 CTCCC
I: FAGG	13 FGA(	1.4 3AA( K	. 14 166/	ST CT
CAC	ATG) C	lcT(	ري 1	ZAA(
36A( D	50 AGA/ E	1410 TGCAC) C T	70 3GCJ A	စ္တ ၃၄ ၁
1290 3CCGG/ P D	1350 GAAGA E E	14] 776( C	1470 GTGG( V A	1530 ACCGC T G
1270 1280 1290 1300 1310 1320 GAGGAGGGATGTCTCTCCAACATGCCGGACACTAGGACGCCGGAGGT G G M C L S N M P D T R T L Y G G R R C	1340 1350 1360 1370 1380 ACCTGGAAGACGAATGTGACTGTGGAGGAATGTA L E D G E E C D C G E E E C K	1390 1400 1410 1420 1430 1440 AGAACCCTTGCAATGCCTCCAACTGCAGGGGGGGGGGGG	1450 1460 1470 1480 1490 1500 GTTCCTGCTGCCACCAGTGCAGGTTCC S C C H Q C K L V A P G T Q C R E Q V R	1510 1520 1530 1540 1550 1560 GGCAATGTGACCTCCCCGAGTTCTGCACGCCAAGTCTCCCCACTGCCCCACCTATT Q C D L P E F C T G K S P H C P T N Y Y
N N	) NGA(	) S	XAAC K	) itt( F
1280 CTCC/ S P	1340 GGAAC E	1400 TGCC7	1460 GTGCA C K	1520 CGAG1 E
L L	رين	N N	C SC O	, CC 1
ire) C	TA(	). C	CAC H	L CTC
1270 GGATC	1330 ACGGG	1390 CTTGC	50 7760 C	1510 GTGAC
12 666 6	13 AAC N	13 CCT P	1450 TGCTGC C C	15 TGT C
GGA G	1330 GTGGCAACGGGT/ G N G Y	AACCCTTGCT	1450 TCCTGCTGCC S C C H	1510 CAATGTGACC
₹5	GI	AG	GT	55

520	540		580.	009
1620 SATGTGCC M C L	1680 SCTČGATC L D L	1740 GCTTGA	1800 TGCCAGA C Q S	1860 ACCTTGA T L N
1610 CTACAACGG( Y N G	1670 ECGGCCTGCC R P A	1730 CTGTGGCAA( C G K	1790 SAAGATTCAC K I Q	1850 CACCACCATO T T I
1600 GGCCTACTGO A Y C	1660 ACCTGGAGC	1720 CTATGGAAA( Y G N	1780 CAAGTGTGGS K C G	1840 ATCTATTGA( S I D
1590 1GGTGGCCA G G Q	1650 \GCTGTGGGG L \ G	1710 TGGTGACAC G D T	1770 CAGGGATGC R D A	1830 1840 CAACGCAGTATCTAT
1580 1590 1600 1610 1620 ACCCCCTGCGAGGGTGGCCAGGCCTACAACGGCATGTGCC T P C E G G Q A Y C Y N G M C L	1640 1650 1660 1670 1680 AACAGTGCCAGCTGTGGGGACCTGGCCTGCCTCGATC Q C Q Q L W G P G A R P A L D L	1700 1710 1720 1730 1740 GEGAATGCTGGTGACACCTATGGAAACTGTGGCAAGGGCTTGA V N A A G D T Y G N C G K G L N	1760 1770 1780 1790 1800 AAGTGCAGTCCCAGGATGCCAGA K C S P R D A K C G K I Q C Q S	1820 1830 1840 1850 1860 CCCCTGGAATCCAACGCAGTATCTATTGACACCACCATCACCTTGA P L E S N A V S I D T T I T L N
FIG. 3f 1570 1580 1590 1600 1610 1620 ATCAGATGGACCCCCTGCGAGGGTGGCCAGGCCTACACGGCATGTGCC Q M D G T P C E G G Q A Y C Y N G M C L	1630 TCACTTACCAGGAA	1690 1700 1710 1720 1730 1740 TTTGCTTTGAGGGGTGAATGCTGCTGGGGCTTGA C F E R V N A A G D T Y G N C G K G L N	1750 1760 1770 1780 1790 1800 ATGGCCAATACAGGAAGTCCCAGGGATGCCAAGTGTGGSAAGATTCAGTGCCAGA G Q Y R K C S P R D A K C G K I Q C Q S	1810 1820 1830 1840 1850 1860 GCACCCAGGCCCCTGGAATCCAACGCAGTATCTATTGACACCACCATCACCTTGA T Q A R P L E S N A V S I D T T I T L N
FIG ATCA 0	TCAC	TTTG	ATGG G	GCAC

5	D
~	)
٢	)
H	-

620	640		680	700
1920 GGAAG E G	1980 CAACC N H	2040 TGGGA G K	2100 TGGCT G W	2160 TTTGC L P
1880 1890 1900 1910 1920 ATCCACTGTCGGGCACCCACGTCTACCGGGGGGGGAGGAAG I H C R G T H V Y R G P E E E G	1970 GTGTGGCCA C G H	2030 GGAAGGCTG	2060 2070 2080 2090 2100 GGCCACGGGGTCTGCAACAACAAGAACTGTCATTGCTTCCCTGGCT GHGVCNNKNCHCFF	2150 CAGTGGTCC S G P
1900 ACCGGGGTCC R G P	1960 CTGGAACCAA G T K	2020 TCTTTGAGAC	2080 AGAACTGTCA N C H	2140 GCAGCGTCGA S V D
DACGTĊTAC IVY	) 19 STGATGACT	) 20 ACCTCCTTC F. S. F	ACAACAAG	ATGGTGGC
1890 3GGCACCC/ G T H	1950 SAGGGCTGG1 G L V	2010 SCAGGAACA( R N T	2070 CCTGCAACAA	2130 CCCGGGAGA P G D
1880 CCACTGTCC H C R	1940 GCTGGACCO	2000 .GGGGCAGTC G Q C	2060 CCACGGGG1 H G V	2120 CTGTAACA( C N T
1870 ACGGGAGGCGGAT G R R I	1930 1940 1950 1960 1970 1980 GGGAAGGTGACATGCTGGAACCAAGTGTGGCCACAACC E G D M L D P G L V M T G T K C G H N H	1990 2000 2010 2020 2030 2040 ATATITGCTTCGAGGGGCAGTGCAGGAACACCTCCTTCTTTGAGACGGAAGGCTGTGGGA I C F E G Q C R N T S F F E T E G C G K	2050 2060 2070 2080 2090 2100 AAAAGTGCAATGGCCACGGGGTCTGCAACAACAAGAACTGTCATTGCTTCCCTGGCT K C N G H G V C N N K N C H C F P G W	2110 2120 2130 2140 2150 2160 GGTCTCCACCTTTCTGTAACACCCCGGGAGATGGTGGTGGCAGCGTCGACAGTGGTCCTTTGC S P P F C N T P G D G G S V D S G P L P
ACC	999 H	ATA	AAA	<b>S</b> .

FIG. 3h

20 AG V 720	80 rc L 740	40 AA T 760	00 TA
2220 rgttggcag L·AV	2280 S A L	2340 3TGGTGGAA G G T	2400 AGGTGACTA
2210 ICTCTTCGT L F V	2270 GGGCAAAC( G K P	2330 ATCTCAGA( S Q S	2390 CAAGCGAAAG
2200 GTTTTCAGC F S A	2260 CCACAAACTGG H K L G	2320 CTTCAGGGTAT F R V S	2380 CCCCAGGG
2190 TCGCTGGGGT A G V	2250 ACAGACAGAG R Q S	2310 TCAGTTGTCC S C P	2370 AGTTGCAGA(
2180 2190 2200 2210 2220 GGGTCCCGTGATCGCTGGGGTGTTTCAGCTCTCTTCGTGTTGGCAG G P V I A G V F S A L F V L A V	2240 2250 2260 2270 2280 cretectacaccacacaccccccccccccccccccccccc	2300 2310 2320 2330 2340 rGCGGCATCAGTTCAGTTGTCCCTTCAGGGTATCTCAGAGTGGTGGAA	2360 CCCAACTITCA
2170 2180 2190 2200 2210 2220 CCCCTAAGAGTGTGCCCGTGATCGCCTGGTTTGGCAG P K S V G P V I A G V F S A L F V L A V	2230 2240 2250 2260 2270 2280 TTCTGGTGCTACTGCTACAGACAGGCCACAAACTGGGCAAACCTCGGCTC L V L L C H C Y R Q S H K L G K P S A L	2290 2300 2310 2320 2330 2340 TCCCTTTCAAGCTGCGGCATCAGTTCAGTGGTGGTGGTGGAA P F K L R H Q F S C P F R V S Q S G G T	2350 2360 2370 2380 2390 2400 CTGGCCATGCCAACCCAAGTTGCAGACCCCCCAGGCAAGCGAAAGGTGACTA

Z

**×** 

24

5

**a** .

۵,

đ

\_\_

F K

**.** 

لم

Z

ď

H

G

٠,_	4
~	)
	٠
C	)
_	4
Ц	4

820	840	. 098	880	006
2520 AGCT S S	2580 AGCC S R	2640 CCAC P P	00 AG	09 . P
25 GAG S	25 AAG S	26. ACC P	2700 AGGAG G G	2760 AGTGC V P
TĠĠ. G	)   	rcg, R	3TC, S	AGC.
O TGC A	JCC P	SCC.	CAG	rccz P
2480 2490 2500 2510 2520 GCCACCTGCACCATGGGCAGCTGCTGGGGAGCT P P A P L S A H L N R A G S S	2540 2550 2560 2570 2580 GCTCGAATAGAAAGGAGGACCCAGGGCCTCCCCCAAGCC A R I E R K E S A R R P P S R	2600 2610 2620 2630 2640 TGCACCTAACTGCCTACTCCCAGGCCTCGACCAC APNCLLSQDFSRPRPP	2660 2670 2680 2690 2700 CCCAGCCAATCCGGTGCCAGGCCAAAGGACCGGTCCCAGGTCAGGAG P A N P V P G Q R T G P R S G G	2720 2730 2740 2750 2760 TCAGCCCCCTACTTCTGGTCCTCCAGCCCCCCAGGCAGTGC Q P P T S G P Q P P R P P A V P
CAG	GAG R	CTC S		CAG.
GAA	CAG R	CTT( F	3AC( T	) ) 
2500 ATCT	2560 CAGC( A	2620 AGGA(	2680 AAAG(	2740 AGCC(
ACA,	STC. S	CA(	SCA.	72 CAC
GGC,	GGA( E	GTC( S	999 <b>\</b>	[CC]
90 GTC S	50 AAA( K	ACT L	70 70 70 70	<u>စ</u> ွင့် ဗ
2490 ATTGT L S	2550 AAGAA/ R K	2610 3CTAC) L L	2670 3GTGC( V P	2730 rTCTG( S G
ACC.	AGA. E	7.TG(	) P	ĨĀĊJ Ţ
O FGC	) VAT/ I	Z Z Z Z	SAA7 N	
2480 ACCT	2540 TCGA	2600 ACCT	2660 AGCC/	2720 GCCCC
	A SGC	GCA A	ر الم	CAG
VTC( S	၁၁	ان 1	וכדל ר	ורי די
2470 TTGAA	2530 AAGCT	2590 TGCCC	2650 AGGCA	2710 CCCTG
26.77 V	2E IGAA E	25 ATG M	26 AAG K	27 TCC S
2470 2480 2490 2500 2510 2520 TGCGCGTTGAATCGCCACCTGCACATCTGAACAGGGCTGCTGGGAGCT R V E S P P A P L S A H L N R A A G S S	2530 2540 2550 2560 2570 2580 CCCCAGAAGCTCGAATAGAAAGAAAGGAGCCCAGGAGGCCTCCCCCAAGCC P E A G A R I E R K E S A R R P P P S R	2590 2600 2610 2620 2630 2640 GACCCATGCCCTGCACTGCCTACTGTCCCAGGACTTCTCCAGGCCTCGACCAC P M P P A P N C L L S Q D F S R P R P P	2650 CTCAGAAGGCACT Q K A L	2710 GCACCTCCCTGCT T S L L

20	GA	H	2880	ည
2820	CAA	×	28	GGT
	<b>AGCTACCCGAGTACCGATCACAGAGGGTTGGAGCAATAATTAGCTCCAAGA</b>	LPEYRSQRVGAIISSKI		GAA
0	TAG	S	0	ATG
2810	AAT	H	2870	ည
	AAT	<b> </b>		GAT
_	AGC	A		9
2800	TGG	ප	2860	ACT
~	GGT	>	2	AAG
	GAG	æ		766
2790	ACA	0	2850	XCGA
21	ATC	S	88	TTC
	2	2		TTC
9	GTA	X	<u>0</u>	TTC
2780	253	B	2840	IAGI
	ACC	Δ.		AG/
_		<b>-</b>	· <b>:</b>	TCC
2770	AAA	×	2830	GTO
~	TCC	<u>م</u>		GAA
	CTGTTCCAAA	>		<b>ICTAGAAGTGTCGAGAAGTTTCTTGTTCCGATGGAAGACTCCGGATGCCATGGAAGGTCC</b>
				~

FIG. 4a

<b>86</b>	78		28	38	-	18	
360 ACG	ATA I	300	AAG A		180	TGC	120
٠٠ کِرَ <b>٢</b>	) H		) E	T T	Š	) 	
CAC	ACA Q		AA( R	Ž o	Š	GT( ₩	
O Y	AAA K	. 067	AAG R	§ æ	0	AAG R	0
350 \GTT7 V		29	KGA E	1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1	170	KCT L	110
GTA V	CAN O		TTATGAAATTATTACTCCTTGGAGATTAACTAGAGAAGAAGGAAG	L G P V L E A G R P D L E Q T V		GCCGCGCGCTCTCGCCCTTGCCTCTGCGACTAAGGTGGCTGC  P R A L S P L A S L R L R W L L	
TT. F	A A	•	JAC.	ر با	Ç	I L	
340 TGAT D	CAG Q	280	VTTA L	25 x	160	S S	100
SAA7 N	ATC I		3AG/ R	, s	, ,	rgc(	
VCCJ	SGTC V		rtG( ₩	A A	2	CCT	
330 TTTT/	rta( Y	270	7 P	E	150	) P	90
SCT.	S	7	IAC T	֓֞֟֓֓֓֓֓֓֟֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓	15	CTC S	90
AGA( D	SAT( I		IAT I		Į.	fi L	
CAC.	3CA O	0	AAT I	کا ہے ۔ ا		CGC A	0
320 \AAC	O CA	260	IGA E	5 1 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5	140	ეე ჯ	80
AG/	S S	•	TTAT Y	<u> </u>	رشن.	) () ()	
E GA	S S		TCI S	7 7	, Abelia	ပ္ပံ့ ပ	
310 CTTG L	ည က	250	TCI	G G	130	r <u>atg</u> gg M g	20
CAC H	99 9	63	CT.	5 J	ר ל	ົວ	
310 320 340 360 360 360 TATTCACTTGGAAAGAACACAGACCTTTTACCTAATGATTTTGTAGTTTACACCTACG I H L E R N T D L L P N D F V V Y T Y D	CTCTGGGGCCCAGGTCCACTCTTACGTCATCCAGGCAAACAGCATA  L G P S S Q Q I S Y V I Q A Q G K Q H I		TCCATCTTCTTCTTATGAAATTATTACTCCTTGGAGATTAACTAGAGAAAGAA	A C G L L G P V L E A G R P D L E Q T V		CCCTCGCTA <u>ig</u> GGGCCGCGCGCTCTCGCCCTTGCCTCTGCGACTAAGGTGGCTGC	

178	Ç •	158			138			118		
TCT S 660 TCA	600 TCT	Œ,	ر ت	540	ب	ÄĞ	480	r P	420	•
AGT V 6 G TGT V	e Agt	H	ACA	വ	~	CAG	4	ਤੂਂ ≈	4 5	
GGGGGAGG	766	S	CTC		7	ACT		<b>y</b> > .	٤	
ATG TGG	590 . AGATG	S	CAG	0	Ç	TGG	0	<u> </u>	ر 0 ا	
GAGA' R ( 650 GCAT( H )	59 GAG	Z	CAA	530	( <u>T</u> ,	CTT	470	<u>ဒီ</u> ပ	410 rrcz	
L L GGA(	ICT	Ħ	GCA		ပ	CTG		₹ 3	ۆ	
P P P P P P P P P P P P P P P P P P P	ည္ဟင္ပ	7	CT		A.	SS		S	ر د	
NTTTACCCCATGGATGCCATCCACGGAGCCTCTGAGATGTGGAGTCT F Y P M D G I H Q E P L R C G V S 620 630 640 650 660 CACAGAGGAAGGCACACAGGGGGGGATGAGGAGCATCCGAGTGTCA T E K E G T Q G D E E H P S V T	580 GGA(	<u>م</u>	SCC	520	S	3AG	460	S L L S D H P N V Q S H C H Y R G	400	
D CAC	Ç.	田	[GA	4,	>	rg.	•	ج ام	ءَ ک	
SCAC H GGGGG	SCA(	·	AAT		A	55		ž z	7	
GCATO GSO CACACO	570 GCAT(	ර	7991	510	>	GT	450	<u> </u>	390	
7666 6 7AC.		[L	IT	ວ	4	Ç	4	₹ ₩	جُ جُ	
SGA7 D GGGGGG	GA7	S	JAG]		S	ľľĆ		2	\ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \	
M M SGA/	ATC	¥	ည		Z	AA1		ر ا	ر م	
P P P P P P P P P P P P P P P P P P P	560 560	z	JAA1	500	ð	SCAC	440	Ž	380 7 7	
TTA( Y NGA(	rta(	四	3GA(		>	\GT		, 1	ۇ ك	
F F YACU	TT		LTT(		G	7995		S S	7	
GAGCACATATTTACCCCATGGATGGCATCCACCAGGAGCCTCTGAGATGTGGAGTCT  E H I F Y P M D G I H Q E P L R C G V S  610 620 630 640 650 660  'AACAGGGACACAGAGGAAGGCACACACAGGGGGGATGAGGAGGAGGAAGGCACACAGGGGGGATGAGGAGGAGGAGGGAG	550 CATA	H	SCA1	490	YVEGVQNSAVAVSACFGLRG	36AC	430	K E G S L L S D H P N V Q S H C H Y R G	370	
GAGCACAT E H I 610 AACAGGGA N R D	Z S	7	S S S	4.	>	GT(	7.	हु स् <u>व</u>	7	2
TTGAGCACATATTTTACCCCATGGATGGCATCCACCAGGAGCCTCTGAGATGTGGAGTCT  E H I F Y P M D G I H Q E P L R C G V S  610 620 630 640 650 660  CTAACAGGGACACAGAGGAAGGCACACACAGGGGGGATGAGGAGGAGGAGGAGGAAGGCACACAGGGGGGATGAGGAGGAGGAGGAGGAGGAGGAAGGCACACAGGGGGGGG	TIGAGCACATATITIACCCCATGGATGGCATCCACCAGGAGCCTCTGAGATGTGGAGTCT	L L H L E N A S F G I E P L H N S S H F	GCTTGCTGCATTTGGAGAATGCCAGTTTTGGAATTGAACCTCTGCACACAGCTCACACT		<b>&gt;</b>	GCTATGTGGAGGGAGTGCAGAATTCCGCGGTTGCTGAGCGCCCTGCTTTGGACTCAGAG		X X	370	110.5

4c
G
FI

218	238	258	278
720 CTGTTCA L F I	780 AGAGAAG R E E	840 CGAATTG	900 3GAGCTG 3 A G
710 ATGTGGAG	770 CTGCTGTGA A V I	830 TAAACATTO N I I	890 TAATTGGA( I G (
700	760	820	880
AGACCCGCT	CGGAACCAGA	ACATCATGT	CTATCAATA
T R Y	R N Q T	I M L	I N I
680 690 700 710 720	730 740 750 760 770 780	790 800 810 820 830 840	860 870 880 890 900
CGCAGAAGAGGCTGTTCTACCACAGACCCGCTATGTGGAGCTGTTCA	GTTGTAGACAAAGGTACGACATGAGGACGGACCAGACTGCTGTGAGAAG	SATGATTCGCTFAGCAAGCATGCATCATCATCGAATTG	CTAGAAATTTGGACAGAGAAATCCTATCAATATAATTGGAGGAGCTG
R R R A V L P Q T R Y V E L F I	V V D K E R Y D M M G R N Q T A V R E E	M I R L A N Y L D S M Y I M L N I R I V	L E I W T D R N P I N I I G G A G
680	740	800	860
AGÄAGAGCTG	AGGTACGACA	AACTACCTGG	ATTTGGACAG
R R A V	R Y D M	N Y L D	I W T D
6	7.	8	80
GCGCAGAA	CAAGGAAA(	CTFAGCAA	ACTAGAAA:
R R R	K E R	L A N	L E I
670 680 690 700 710 720	730 740 750 760 770 780	790 800 810 820 830 840	850 860 870 880 890 900 IGCTGGTTGGACAAATTTGGACGACAGAAATCCTATCAATATAATTGGAGGAGCTG
CTCAGCTGCTGCAGAGAGAGAGCTGTTCTACCACAGACCGCTATGTGGAGCTGTTCA	TTGTTGTAGACAAGGTACGACATGGGACGGAACCAGACTGCTGTGAGAAGA	AGATGATTCGCTFAGCAAGCTAGCATCGTTAAACATTCGAATTG	
Q L L R R R R V L P Q T R Y V E L F I	V V D K E R Y D M M G R N Q T A V R E E	M I R L A N Y L D S M Y I M L N I R I V	
0	. <b>⊢</b>	<b>∀</b>	<b>.</b>

Œ

318	338	358	378	308
980 990 1000 1010 1020 CAGTTGGTTTTGAAGAAAGGCTTTGGTGGAACTGCCGTTTG 0 L V L K K G F G G T A G M A F V	1030 1040 1050 1060 1070 1080 TAGGAACAGTATGTTCAAGGAGCCACGCAGGTGGGATCAATGTTTGGGCAAATCACTG G T V C S R S H A G G I N V F G Q I T V	1090 1100 1110 1120 1130 1140 TGGAGACATTTGCATTGCTCATGAATTGGGGCATAACCTTGGAATGAAT	1150 1160 1170 1180 1190 1200 ATGATGGGAGAGTGTTTCTGTGGAGCAAAGAGCTGTATCATGAATTCAGGAGCATCG D G R E C F C G A K S C I M N S G A S G	1210 1220 1230 1240 1250 1260 GGTCCAGAACTTTAGCAGTTGCAGTGCGAGGACTTTGAGAAGTTAACGTTGAATAAGG S R N F S S C S A E D F E K L T L N K G
TGGC	AAAT	TGAA'	GAGC. A	TGAA) N
1010 AGGAA G M	1070 TGGGC G Q	1130 TGGAA G M	1190 TTCAGG S G	1250 AACGT T L
CTGC. A	rett.	ACCT	rgaa? N	\GTTZ
00 GGAA( G T	60 AATG	20 CATA/ H N	1180 GTATCAT I M	10 3AGA/ 3 K
1000 TGGTGG/ G G	1060 GATCAA: I N	1120 GGGGCA7 G H	118 CTGT/ C	1240 CTTTGA( F E
GCTT	GTGG	AATT! L	AGAGI S	AGGA( D
990 JAAAGGC K	1050 CGCAG	1110 CATGA H E	1170 GCAAA A K	1230 GCGGA A E
rgaac K	CCA(	TGCI	TGGA	CAGT S
980 GTTT1 V L	1040 AAGGAG R S	1100 :CATTG1 I V	1160 . TTTCTG7 F C	1220 CAGTTG S C
GTTG	TTCA S ]	ATCCA	TETE C I	TAGC/ S S
ACA O	'ATG' C	TGC.	AGA(	CTT
970 GTGC/ A	1030 CAGT/	1090 ACATTT T F	1150 GGAGA	1210 GAAA(
ACGACAGTGCA	1030 IGGAACAGTI G T V	GAGA E T	1150 GATGGGAGA( D G R I	TCCA S R
AC	TA	TG	AT.	99

<b>4</b> e
4
G
丘

•	418	438	458	478	498
1320 TCCTGTG	S C G	1380 TGTGAGG C E V	1440 GCATATG A Y G	1500 AAGACCA K T S	1560 GATGTCT D V F
1310 AGCGCGCC	S A P	1370 GCGAAGGAG A K E	1430 GCTGAGTGT A E C	1490 TGCAGAGGG C R G	1550 TGCCCGCCA C P P
1300 GAAGCCTAC	E A Y	1360 TGCGGCACA C G T	1420 AAGTCATTT K S F	1480 GGCTCCATG G S M	1540 TCTCAGTTC S Q F
1290 GAAGCCTGA(	LLNIPKPDEAYSAPSCG	1340 1350 1360 1370 1380 GTGGACCCTGGAGGAGTGTGACGG V D P G E E C D C G T A K E C E V	1400 1410 1420 1430 1440 TGFGAAGGAACTTGTAAGCTCAATTTGCTGAGTGTGCATATG C E G S T C K L K S F A E C A Y G	1460 1470 1480 1490 1500 AAAGATTGCCAGTTCCTTCCAGGGCAGGCAGACACA K D C Q F L P G G S M C R G K T S	1520 1530 1540 1550 1560 GTTCCTGAGTACTGCTTCCTCTCAGTTCTCT V P E Y C N G S S Q F C P P D V F
1280 TTAACATCC	N I P	1340 ACCCTGGAGA	1400 AAGGAAGCAC GST	1460 ATTGCCAGTT C Q F	1520 CTGAGTACTG E Y C
1270 AGGAAGCTGCCTGC	7 7 2 8 9	1330 TAATAAGCTGGTGG N K L V D	1390 GGACCCATGCTGFG D P C C E	1450 CGACTGTTGTAAAG D C C K D	1510 FGAGTGTGTTC E C D V P
CTGCT	G S C L L N	1330 1340 1350 1360 1370 1380 GTAATAAGCTGGTGGACCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGA	ن ت	<b>~</b> ~	O.T.

<b>4</b> £	<b>→</b>
C	)
互	_

	518	538	558	578	598
1620	c Q	1680 AAGAG R D	1740 :ceeca e s	1800 GAATG N V	1860 AGGCA G T
1610	ATGGCAT G M	70 STGCCCC A P	30 F. S. F. S.	90 VATGCGA C E	50 CCAGTCG S R
16	GCTACA	1670 CTAAGGCTC K A A	1730 ACTGTGGT1 C G F	1790 AGCTTCAA1 L Q (	1850 AGACACCC
1600	GCCTACT A Y C	1660 TCAAAGG S K A	1720 TTTGGCA F G N	1780 TGTGGAA. C G K	1840 ATCATTC I I 0
1590	TCATTCAGAATGGATATCCTTGCCAGAACAGCCTACTGCTACAATGGCATGTGCC I Q N G Y P C Q N S K A Y C Y N G M C Q	1630 1640 1650 1660 1670 1680 AATATTATGACGCGCAGTGTCATCTTTGGTTCAAAGGCTAAGGCTGCCCCAAGAG Y Y D A Q C Q V I F G S K A K A A P R D	1690 1700 1710 1720 1730 1740 ATTGCTTCATTGAAAGGTGACAGATTTGGCAACTGTGGTTTCTCCGGCA C F I E V N S K G D R F G N C G F S G S	1750 1760 1770 1780 1790 1800 GTGAGTACAAGAGCTGCGAACGCCTGTGGGAAAGCTTCAATGCGAGAATG E Y K K C A T G N A L C G K L Q C E N V	1810 1820 1830 1840 1850 1860 TACAGGACATGCCGGTGTTTGGAATAGTACCAGCTATCATTCAGACACCCCAGTCGAGGCA
<b>H</b>	GCCAGA Q N	10 AGGTCA V I	1. CTAAAG( K G	TTGGGA/ G N	18 GAATAG1 I V
1580	ATCCTT PC	1640 AGTGTC	1700 TCAATT	1760 GTGCCA( A T	1820 TGTTTG( F G
0	ATGGAT G Y	O ACGCGC A Q	0 TTGAAG E V	O AGAAGT K C	O TGCCGG P V
1570	ATTCAGA I Q N	1630 AATATTATGAC Y Y D	1690 TGCTTCAT1 C F I	1750 GAGTACAAG E Y K	1810 CAGGACATG Q D M
	TCA	AAT	ATT	GTG,	TAC

ģ
4
G
L

618	638	658	819	869
1890 1900 1910 1920	1950 1960 1970 1980	2010 2020 2030 2040	2070 2080 2090 2100	2130 2140 2150 2160
FTCCAGCTTCCGACGTTCCAGGGATGGTGA	3CTGGCAAGATTTTTCAGTGTGTAAATGCTT	SACATTCAGGGAAAATGTCATGGCCATGGGGTATGTAACA	SAAGATGGCTCCCCCACACTGTGACAAGGAT	GCCGACGTATAATGCAAAGAGCACACACACGACG
F Q L G S D V P D P G M V N	1 G K I C R N F Q C V N A S	O I Q G K C H G H G V C N S	S D G W A P P H C D T K G Y	P T Y N A K S T A L R D G
1880	1940	2000	2060	2120
TGTGGAT7	ATGTGATC	reacteta	CACTGT(	GACAGCG
V D I	C D /	d c i	H C 1	D S G
1870	1930	1990	2050	ATGGAGGAAGCGTG
CCAAATGCTGGGGT	ATGAAGGCACCAAA	CTGTCCTGAATTAT	GCAATAAGAATTGT	
K C W G	E G T K	V L N Y	N K N C	

Z

G

H

œ

S

Ç

G

9

 $\alpha$ 

<b>4</b> h	
<u>ن</u>	
Ĭ	

738	758	778	
2280	2340	2400	2460
ATGGCA	CAGGGG	GCCCAG	
G R	G G	P G	
2270 CAAATGTCAG Q M S D	2330 FCCAGACCAC	2390 CACCAGGGG	2450
2260	2320	2380	2440
GAAGAGATCA(	TCCTAGTATCT	IGTCTCCAGAC	
K R S (	P S I S	V S R P	
2250	2310	2370	2430
CCTTCAGGAA	AGCCAGGAGA'	3GGCCCAGG	
F R K	P G D	G P G	
2240 TACGGAAAA( . R K T	2300 TCTCTAGACA	2360 GACCACCAGO PPG	2420
2230	2290	2350	2410
AAGAGAGATGAAC	AATCAAGCAAACG	CCAAATGTCTCCA	
K R D E I	N Q A N V	P N V S R	
	2240 2250 2260 2270 2280 GAACTACGGAAAACCTTCAGGAAGAGAGATCACAAATGTCAGATGGCA E L R K T F R K K R S Q M S D G R	2240 2250 2260 2270 2280 GAACTACGGAAACCTTCAGGAGAGAGATCACAATGTCAGATGGCA E L R K T F R K R S Q M S D G R 2300 2310 2320 2330 2340 AACGTCTCTAGACAGCCAGGAGATCCTAGTATCTCCAGACCACGGGG	2240 2250 2260 2270 2280  3 L R K T F R K R S Q M S D G R  2300 2310 2320 2330 2340  ACGTCTCTAGACAGCCAGGAGATCCTAGTATCTCCAGACCACGGGG  V S R Q P G D P S I S R P P G G  2360 2370 2380 2400  CCAGACCACGGGCCCAGGTCTCTCCAGACCACGGGGCCCAG  3 R Q P G D P S I S R P P G G  3 R P P G G P G V S R P P G G P G

838

818

FIG. 4i

	7	2530			2540	2540 2550 2560 2570 2580		25	2550		8	2560			2570	0		2580	80	
CACCACAACGAAAATTCTCTCAGGGAAACTTGATTCCGGCTCGGCCCGCTCCTGCAC  P Q P K I S S Q G N L I P A R P A P A P	ACA	ACC.	GAA	AAT	ATC	S	TCA 0	999	AAAN	CTT	GAT	TCC	GGC	TCG R	ည်	ည် (၁၅၃	JCC P	TGC	AC P	•
2590 2600 2610 2620 2630 2640	2	2590	ر پر	ريل	2600	0,5	ر و ز	26	2610		Ö	2620	Ė		2630	0		2640	40	

 $\vdash$ 

S

FIG. 5a GCCAGAGTAG CGCGCGCGCG TTTTTTGAA AAAATGAAAG TGCGCGAGGG GGTCGCGGCA
FIG. 5a  GCCAGAGTAG CGCGCGCG CACGCACACA CACGGGGAGG GGAGAAAGTT TTTTTTGAA AAAATGAAAG GCTAGACTCG CTGCTCAGCG ACCCGGGCGC TGCGCGAGG GGTCGCGGCA GACTCAGGGC AGTAGGACTT CCCCCAGCTC GGCGCCCGC TGGATGCTG CACCGCTGC CGCGGGCCC CCGAAGCAGC TGCACGCCGC TGGATGTG CACCGCTGC CGCGGGCCC CCGAAGCAGC TGCAGCCCGG GCCGCGCACA ATGGCAGAGC GCCGGCGCC CCGAAGCAGC TGCAGCCCGG CCTCCTGCT GGCCCTGGCT GGGCCCTGC TACGAAGTGG CCAGAGCCTC CCTTCTGAGC AAGGACCCTG GGATCCCAGG ACAGAGCATC CCAGAGCCTC CCTTCTGAGC AAGGACCTG GGATCCCAGG ACAGAGCATG CCGAGACCTG ATCCTCAGC TGGAAAGGAA TGAGGGACTC ATTGCCAATG GCTTCACGGA GACCCATTAT CTGCAAGATG GTACTGATGT CTCTCCACT
FIG. 5a  GCCAGAGTAG CGCGCGCGCG TTTTTTGAA AAATGAAAG TGCGCGGGGGCGCA TGCGCCGCGCGCGCA TGCACGCCGCG TGGGATGCTG TGCACGCCCGG CCCTCCTGCT TGCAGCCCGA GGGATGAGTT TGCAGCCCGA GGGATGAGTT TGCAGGCCTC CCTTCTGAGC CCAGAGCCTC CCTTCTGAGC TGCAGGCCTC CCTTCTGAGC TGCAGGCCTC CCTTCTGAGC TGCAGGCCTC CCTTCTGAGC TGCAGGCCTC CCTTCTGAGC TGCAGGCCTC CCTTCTGAGC TGCAGGCCTC ATCCTCAGGC
_ , , = =

1050

1100

1150

1250

200

300

350

400

AGCAGCTGCA

CATGGTGTTC

TCCCATTCCC

TCCACGGGGT

CATGAACCCG

1000

750

FIG. 5b

800

850

900

950

#### TGTTTGAAAA TAAAACGTAC AGCTTAGAGC CAATGAAAAA CACCACTGAC AAGGGCTGTG TGGGTCACAG CATAACAAGT CCAACCTCAC CATGGAAGAT GTCTCCCCTG GAACCTCTCA AATGCGGGCA AGAAGGCATA AGAGAGAGAC CCTTAAGATG TAGAGCTGGT TATTGTGGCA GACAACAGAG AGTTTCAGAG GGAGTGGAAG TGTGGAATGA CATCGACAAA TGCTCTATAA GCCAGGACCC GAGATCGCCA CGTGCTGGTA GAGAAGATA AAGCTTCTAC TTATTTCCAA CTGCAGAACA GTCTGGAGGA GTTGTCATGG ACCATTCAGA CAGCCCCCTT GGTGCCGCAG TGACCTTGGC ACATGAGCTG GGCCACAACT TCGGGATGAA CCATGACACA GCTGCAGCTG CAGAATGGCC GCAGAGAAAG GAGGCTGCAT TCAGTGGGGT TCGTCCCAGC TGAGAGCATG ACGAACATCC GCGATTAATA ATCACGTTGA CAAGTTTTAC AGACCACTGA ACATCCGGAT AGCATGTGCA CTCCATGAGT TTCTAGACTG AAGTTAAGCA CCACGACAAT GCTCAGCTTA GGAACCACCA TCGGCATGGC ACCCATCATG GACCTGGAGA ACCAAGTACG ATTCACCAGG AGCTACAAAC GCAAGGAAAA CTCGAAAATC CTGGAGAGGG

110.00					
GCAGGAAGGA C	CCTGGAGGCT	AGCCTGGAGA	CTGGAGGCT AGCCTGGAGA AGGGCATGGG GATGTGCCTC	GATGTGCCTC	1450
TTCAACCTAC	TTCAACCTAC CAGAGGTCAA GCAGGCCTTT	GCAGGCCTTT	GGGGCCGGA	AGTGTGGAAA	1500
TGGCTATGTG G	GAAGAGGGAG	AAGAGTGTGA	CTGCGGAGAA	CCGGAGGAAT	1550
GCACGAATCG C		TGCTGTAAC GCTACCACCT	GTACTCTGAA	GCCAGATGCT	1600
GTGTGCGCGC A	ACGGGCAGTG	CTGTGAAGAC	CGGGCAGTG CTGTGAGAC TGTCAGCTGA	AGCCTCCAGG	1650
AACTGCATGC A	AGGGCTCCA	GGGGCTCCA GCAACTCCTG TGACCTCCCA	TGACCTCCCA	GAATTCTGCA	1700
CAGGGACTGC C	CCCTCACTGT	CCAGCCAATG	TGTACCTACA	TGATGGCCAC	1750
CCGTGTCAGG	CCGTGTCAGG GCGTGGATGG	TTACTGCTAC AACGGCATCT	AACGGCATCT	GCCAGACCCA	1800
TGAGCAGCAG T	TGTGTCACGC	GTGTCACGC TCTGGGGACC AGGTGCTAAA	AGGTGCTAAA	cceerccre	1850
GCATCTGCTT		TGAGCGAGTC AACTCTGCAG GAGATCCTTA	GAGATCCTTA	TGGTAACTGT	1900
GGCAAÄGACT	CCAAGAGCGC	CTTCGCCAAA	TGTGAGCTGA	GAGATGCCAA	1950
GTGTGGGAAA	¥	TCCAGTGTC AAGGTGGTGC AAGCCGACCT	AAGCCGACCT	GTCATTGGTA	2000
CCAATGCTGT	←	TCCATAGAA ACAAATATCC CACAGCAGGA	CACAGCAGGA	AGGAGGTCGG	2050
ATTCTGTGCC	GGGGGACCCA	TGTGTACTTG	TGTGTACTTG GGTGATGACA	TGCCAGACCC	2100
AGGGCTTGTG	CTTGCAGGAA	CAAAGTGTGC	CTTGCAGGAA CAAAGTGTGC AGAAGGAAAA ATCTGCCTCA	ATCTGCCTCA	2150
					٠

# FIG. 5d

2200	2250	2300	2350	2400	2450	2500	2550	2600	2650	2700	2750	. 2800	2850
GTGTGCCATG	GCCACTGTGA	GGAGGAAGCA	GACTGTAGGA	TGGTGTATCT	AAAACCACCA	TGGCCCTCAC	TGAACCGGGC	TGCCAGAACA	GCTTCAGTCA	CACCCAGTGG	GCCCAGGGCA	TCCACTGAGC	GGCAGCAGGA
ATCGTCGATG TCAGAATATC AGTGTCTTCG GCGTTCACAA GTGTGCCATG	ATGTAACAAC AGGAAGAATT	TCTGTGACAA GTTTGGCTTT	CAAGCAGATA ACCAGGGCTT GACTGTAGGA	GCTGGATTTG TGGTGTATCT	CACACATAAA AAAACCACCA	GGACACCCAG	CTTGGCCAGG CTCACCACAC CCCCGGGAAA GGCCTGCTGA TGAACCGGGC	CTCGCTGAAA	TGGACATCAG CAGGCCCCTC GACGCTCGAG CCGTCCCACA GCTTCAGTCA	TCTCCACCAG ACCCCAGGG	CCTGCCAGG CCCTGCCCG CCAGTCCTGC AGTCAGGCAG GCCCAGGGCA	CAGAAGCCTC TGCCTGCTGA TCCACTGAGC	TGCCTTGGTG AGGACCCCAG GGCAGCAGGA
AGTGTCTTCG	ATGTAACAAC	TCTGTGACAA	CAAGCAGATA	TCTGCTTGCT		CACCCTTCCC	CCCCGGGAAA	AGGACAGGCA	GACGCTCGAG	TCTCCACCAG	CCAGTCCTGC	CAGAAGCCTC	recerreere
TCAGAATATC	CAGTGCCACG GCCGAGGGGT	AGCCCACTGG GCTCCACCCT	CAGACAGIGG ICCCAICAGG	ATCCTGGTGA GCATCCTGTG TCTGCTTGCT	CAAAAGGAAG ACGTTGATGC GGCTGCTT	TGGAAAAGCT AAGGTGTGTG CACCCTTCCC	CTCACCACAC	ACCACATTIC AATACCCCCA AGGACAGGCA	CAGGCCCCTC	recrecce	CCCCTGCCCG	TTCGAAAACC CAGTCCTCCT	GGCTCACTAG
ATCGTCGATG	CAGTGCCACG	AGCCCACTGG	CAGACAGTGG	ATCCTGGTGA	CAAAAGGAAG	TGGAAAAGCT	CTTGGCCAGG	ACCACATTTC	TGGACATCAG	CCTCAGCGAG	CCCTGCCAGG	TTCGAAAACC	AGGACTICIC GGCTCACTAG

READING FRAME	Æ			
AGČCC	CCATCAGACC	ACCTGGGCAC CGCCCAGCCC CCATCAGACC TGCCCCTAAG CATCAAGTAC	CATCAAGTAC	2900
TGCC	TATATCAAGT	CCAGACCTTC CCACAATGCC TATATCAAGT GAGAAGCCAG CCCAGACCGG	CCCAGACCGG	2950
CAGA	AGTTTGCACT	TCCTCAACAG TGAAGACAGA AGTTTGCACT ATCTTCAGCT CCATTGGAGT	CCATTGGAGT	3000
rtcc	GAGTTTCTAA	TGTTGTTGTA CCAACTTTCC GAGTTTCTAA AGTGTTTAAA ACACCATTCT	ACACCATTCT	3050
ACT	GCCATCGGTG	CTCCAGACCC TGGAGCCACT GCCATCGGTG CTGTGTGTG GTGCTTTGTG	GTGCTTTGTG	3100
GTÄ	AGTTATTAAT	TACTTGCTCA GGAACTTGTA AGTTATTAAT TTATGCAGAG	TGTCTATTAC	3150
CAG	GCATTTGTAC	TGCGCAGGGC GCCGTAGCAG GCATTTGTAC CATCACAGGG CTTTTCTACA	CTTTTCTACA	3200
GCT	TTTGTTTTC	GAAGGAAGGC TCCTCGTGCT TTTGTTTTC TGGAGGACTT GAAATACCCT	GAAATACCCT	3250
ATG	AGATGTŢTAC	GCTTGATGGG ACCTAAGATG AGATGTTTAC TTTCTATTCA AGGCCTTATC	AGGCCTTATC	3300
CTL	CCCAAGGCTG	GGAAAATAGC TCCCCACCTT CCCAAGGCTG TTATGGTACC AGACACAG	AGACACACAG	3350
AGA	ACCTGGCATG	CTCAGGACAC CCCAGGGAGA ACCTGGCATG GGTTTTCTTT GTTTGCTTTC	GTTTGCTTTC	3400
TTG	ATTITATCIT TTATATITIG GTATCCCTAT	CTTGGGTTGT AGCCAGGGCC	AGCCAGGGCC	3450
CCA	CIGCATGCIA	TICAGGAAGG ICTIGGGCCA CIGCATGCIA AIGGCCTICA GGICCIGCAC	GGTCCTGCAC	3500

	3550	3600	3650	3700	3750	3800	3850	3900	3950	4000	4050	4100	4150	4200	4250	
			TGAGCTGAGC		TTTGCAGAGT		CCCATTTAGA		ATTTTCANAT	•				ACTCGGTCTT		
	CCTGAAGCTC TCAGACAACA AGTAGGATCT GCTTTCTAGC CAGCAGCTTT	TTGGGGTGTG GTTATACCAG	CATCTGACCT TO	AGCCATGAGC ACCTCTAGGA AGCAAGGACG GCTGAGGTGC TGCACAAGGC	CLGGCTGCCC	GTTTACATCG GGAACAGTGG TGTTTCTACA	ccercrccre co	GTCTTCTGGA CTGTAATACA ATGACCCTGT	AACCATAACT	AGTCGATCAA CTCTGGAGAT AGAAATCTCC	TTATTAAAAA TTAGGCAGAA TCCATATGCT	TCTCATGGCA CAGCCTGAGT	TTAGTAGC CATTGGACAA AGCACCCAAA GTTACCTGTG	GCATAGAGAG	CCTCACATTC TGAACATACC TATCAATGAC TAAGNCAGCA AGGCAATCCG	
	AGTAGGATCT	AAAAGAAGGT	TCTGGGCAAA	AGCAAGGACG	GGGCTTCTCT		CTGCAGACCT	GTCTTCTGGA	CTTTCTATGG	AGTCGATCAA	TTATTAAAAA	GGAATGCTCT	CATTGGACAA	AATTTCTTCA	TATCAATGAC	
	TCAGACAACA	GGAGAGAACC TGGGGTACTG AAAAGAAGGT	GGAATCCTAA	ACCTCTAGGA	GAGCTGGCAG	GCTAGCTGGC ATGGCATGTT	TGCCTGGGCA	TTACCACATT	ATAGAGGAGG	AACCAGATCT	AAGGCTCGAC	ATAACCACGT		AAGGCATCCT	TGAACATACC	
10.)	CCTGAAGCTC	GGAGAGAACC	GATGGAGACT GG	AGCCATGAGC	TCTGCTTTGA GA	GCTAGCTGGC	AGAAAGCCAC TG	GCTAAGCAAA TT	GTTCTGACAG AT	GTGAACTAGT AA	TTTTACTGC AA	TGCAAAAGCT AT	CTGGTATCCT TA	TGTTCTCTTC AA	CCTCACATTC	
				•												

۵ · ۵ ·					
TTTCCGAATA C1	CTGAGTTGCT	rgagttgct cacggnaagg caacctcagc ccaggnaaac	CAACCTCAGC	CCAGGNAAAC	4300
TTTTTCCTC	TTTTTTCCTC TGNTCTTTCA	GTATGTGACT	GGGGAGCTAC CTTCAGAAGC	CTTCAGAAGC	4350
AAATTTTCAA	AAATTTTCAA GGTGGNCTCA ACCCCATNGG	ACCCCATNGG	ATGAAAGNTA	TTTTTTAAA	4400
AAATAATTAA	AAATAATTAA TGGTAATGCC AGAGGGCTTT	AGAGGGCTTT	CCTGGCNTCC AGATNGGGGC	AGATNGGGGC	4450
GTAGGNTTGA	GTAGGNITGA CTAGCIIICA	CGACAGAAGG	TAAATGACAG CAGTCCTCA	CAGTCCTCTA	4500
CCTCGTCTGA	CCTCGTCTGA CTGCTTTAAG ATCAAGGCTT	ATCAAGGCTT	CTTTGGAAGG GTAACTAACA	GTAACTAACA	4550
TTAATGGCTG	TTAATGGCTG GCCTGTGCCT	TGAAGCAGAA	GGGAAATAC AGATAAGGAA	AGATAAGGAA	4600
TTTGGTTTGC	TTTCTAGAAT	CCAAAACTGT	ATCCAGCATT GGGAAGCATG	GGGAAGCATG	4650
GTCTTCATGA	CTGGGTAAAT	AAATCCACGT	CACAGATGCA	TAAAAGAATA	4700
ACTCTTATGA	ACTCTTATGA CATGCCTCTT	TTTGTGGCAC AGAGACAATA	AGAGACAATA	TTGCTGCCAC	4750
TGAGATGCAT	TGAGATGCAT ACAAATTTC TGTAACTGAT ATGTCATTCA	TGTAACTGAT	ATGTCATTCA	GTAGTTGTAT	4800
TAAGGCCAAA	TAAGGCCAAA CATCCACAAC	TGTAAAGACT	TATAGAGTTG	TGTGGGCGTT	4850
GTCTTGTGAG	GICTIGIGAG ACACACAAAG CCICAGCIGA AGCGIAIGAG	CCTCAGCTGA	AGCGTATGAG	crccrccrc	4900
AGGTGGGAGT	GATGGGGAGG		CTAGAACAC ACAAAGACAA	CAGAAGAGCT	4950
TTGGTTTGGG	GGGGGTGCAG AGAGAGTGTG GTTTAGAGGA AGTTGGAGCC	AGAGAGTGTG	GTTTAGAGGA	AGTTGGAGCC	2000

ATGATCTTCT	ATGATCTTCT GCCATCTCCC CAGTGTCCAC TAAGGATGCC GATGGTGCCT	CAGTGTCCAC	TAAGGATGCC	GATGGTGCCT	5050
TACCAGCTGT	GCAGTGCTGG CTGCTTGCTT	CTGCTTGCTT	TTACAGAGCC ATGCATTCAT	ATGCATTCAT	5100
TTCTGAATAA GA	GAACATATTT	AATCCTGAAA		TICCCITACA GGACAGACÄG	5150
TGTTACTAAA	GGAATTCCTC	TAAGATACAG	TAAGATACAG TAGTTGTCAA	TTAAAGCATA	5200
TTTAGCAGTA	ACTTCAATTT	TAACAAATT	GGGACCCAAT	AGCCAGCATG	5250
AGGGTTCTTT	GACAGAGGGT	AGTTTCTCTC	AGTITCICIC ICCCIIICIC CAICCIICAA	CATCCTTCAA	5300
ATGACAAGAC	GTCAAAACTA	ATACAGTTCA		TTTGCAGTCC ATCTCATGCT	5350
TATACATACT AG	AGAGGTATGA	CTAAAGTTGG	CTAAAGTTGG TTGAGTCATG GGAGACCATC	GGAGACCATC	5400
CCTGAGAAAG	TC	CAGTCGGT CAAGAGCCTT		GCCAGGTGGC GTGGCTGGAC	5450
GICCICCIII	TGTTCCTGCA		CTGAGGAATA GTTATAGGTT ATGTGACCCC	ATGTGACCCC	.5500
ACTTCACAGG	CA	GCGAACCTTG	AGTGGGAG GCGAACCTTG CAGGCATGCC CCTTAAAAGC	CCTTAAAAGC	5550
TGGTCTCAGA	ည	TCCTGAGTCT	TACAATAG TCCTGAGTCT GTTTTCCCAG CACAGAGA	CACACAGAGA	2600
GCAACAATGC	AGTTTTCCAT	<b>TTCAAAATAT</b>	TTCAAAATAT GCATGCCGAG TTTGCGCTCT	TITGCGCTCT	2650
GTGTGAGTGT	TTCCAGGTTA	CACATATGGG	CACATATGGG ATGACATCAC	AGAAACCACA	2100
CAAGCAACAA AT	ATTAAATTCT	ACGGGAAGAA	TAAATTCT ACGGGAAGAA ATCCTCCTGA CTGGTCTCTG	CTGGTCTCTG	5750

1)					•
AGGAGACATT	TTTATGCCTT	CTTAACTTTA	ITTATGCCTT CTTAACTTTA TTAGGAACTC TCAGGCTGAA	TCAGGCTGAA	5800
GCTAGGGGTC	ATTGTCCCCC	AACAAATCAA	TTGTCCCCC AACAATCAA TACAAAGCCA TCAATGNACT	<b>TCAATGNACT</b>	5850
CTCGAAGAAC	TGCCAAACCC	GCCAAACCC TGATCTGTGT	GAATGTTCTC AGGAGCCTGT	AGGAGCCTGT	2800
GATCCCCATG	GTGCTANAAA	GAGGCTGGAG	GTGCTANAAA GAGGCTGGAG CTGGGCCAAC AAGAAGGCCT	AAGAAGGCCT	5950
AAGAGTCCTC	CTGCCTCTCA	CTGCCTCTCA GCAGATGTTT	ACTGAGCACT CTGAGCCAGA	CTGAGCCAGA	0009
AGCACCCCGA	CAACCAGGAG	GACGATNGCT	CAACCAGGAG GACGATNGCT GGGCAGTAGG GCGCCCAGCC	GCCCCAGCC	6050
ACTTGCAGCT	CTITCCTCTG	STITCCICTG AGGCCCGCTT	TGTGTTTTAA TTCCCTTCTG	Trccctrcrg	6100
TCAGGCCCCA	ANCAGNGGAC	ACTGTCCTAT	ANCAGNGGAC ACTGTCCTAT AGACCTCCCT CTNAGTTTTC	CTNAGTTTTC	6150
AGACGGCCTA	AGCCATACAC	AAATGCCCCA	AGCCATACAC AAATGCCCCA GACTAAGAAA CACCAATACN	CACCAATACN	6200
TCCCAGCAGT	CCCCAAGAAC	TGGTTTTTAA	TGGTTTTTAA ACACTATGAC AAGTAGAAGA	AAGTAGAAGA	6250
GGGTGTCACA	GAGGCCATTT	TTTTTCTTTT	CTTTCCACTC ATACTGGAAC	ATACTGGAAC	6300
CTAGGTCCTC	TCTCTACACT	CCTAGTTCCT	TTACACAACT	CGGCAGTGGC	6350
TCCATTACAC	CAAGGACACA	GAAAAACACA	CAAGGAÇACA GAAAAACACA GGTACCGATT	TGCCTTCCTC	6400
TCCTGCCAAT	CACAAGTGCC		TTACTCTGAC CAGACCCATG ACAAAACCTC	ACAAAACCTC	6450
TGTCATCCAA	GAGAGCCAAC	TCTCTACCTT	SAGAGCCAAC TCTCTACCTT TGTTACTACT TCAAGCCAAT	TCAAGCCAAT	6500

### 0029 6750 6.650 6800 6850 0099 0069 CATATGAATG TACCATTTCC ATGCCTTTTG TGGAGTACAG ACATATAAAC CACCAGGACC ATAGCACAGA GCAACCTCCA GNACACACAC ACACACACAC CTTGAATCTA TCCCACAGCA TATCAACCCA CAGTGACCTC CCTCCCACCG CCTTGTTCTA ATTACAAGGT GAAGATGGCC ATAGAAAATC AAGTTAGCAC TGCTCAGATC CACTGGGCAG GGGGGACTCC TTGCAGGAGA CATGGGCACA GCTTTGAAAC TCTGATGAGT TAAGTCATGC TCTGGGAGCT GTGAGCCCCA GTGGTAACTG CTAACCTTCA AGGGTCACCT AAACAGTATA GTCCAACCTT TAATTACAAA ATGCTTTTGA TGCAACCTGA ATTTCCCAAT GGCACCTATT ATAAATACTT CCATT FIG. 5j

	09		120	180	.240	300	360	420	480	540	009	099	720	780	
	GCCGTTGCTA		TCATGCCCGG GCGCGGGC GTCGCCCGGT TCTGCTTGCT GGCTCTCGCT CTGCAGCTAC	ATTGGCCGCT GGCGGCGTGC GAGCCGGGAT GGACCACAAG AGGAAGCCAA GAAGGTAGCC	ATAATACCTC AGTGGCGGAC TTCAGAAAGC CCTGGGAGAG	CTGATCCTAG	TACACTGCAA	CACGGGACTG	CCGGGGAATT AGAGGACTGA	AGCTACATCA TCGAGCCCGT CCCTAACAGC GACAGCCAAC	TTCGAGCACT	AAGCAACCTC	CTGGTGGCTG	AAGCTCATGG	
	CAGTGAGAGG		GCTCTCGCT	AGGAAGCCAA	TTCAGAAAGC	AGGGCGAGAG	AGAAACCTGC	CTGCTTTTAC	CCGGGGAATT	CCCTAACAGC	GAACTGTGGG	TCAGACCAAA	GGAGCTTTAC	CACCAAACGC	
	TGATTGCGGA		TCTGCTTGCT	GGACCACAAG	AGTGGCGGAC	TCATGGCTGA	CAGCCTACAC	CTGAGGATCA	TCAGCACCTG	TCGAGCCCGT	TGCCCCGGG	AGTTTACACA	TGAAGTACGT	ACCAGGATGC	
	AGGGGATGTG	READING FRAME	GTCGCCCGGT	GAGCCGGGAT	ATAATACCTC	GAGCTCAGGG TCATGGCTGA AGGGCGAGAG CTGATCCTAG	CTTTTTGCIC CAGCCIACAC AGAAACCIGC TACACTGCAA	ACGCTGAAGT CTGAGGATCA CTGCTTTTAC CACGGGACTG	AGTGTCACGC TCAGCACCTG	AGCTACATCA	CATCTCACGC TGCCCCGGG GAACTGTGGG TTCGAGCACT	TGGGCCCTTC AGTITACACA TCAGACCAAA AAGCAACCTC	CTACACTCTA TGAAGTACGT GGAGCTTTAC CTGGTGGCTG	AATCGACATG ACCAGGATGC CACCAAACGC AAGCTCATGG	
	GECCEGEGEC AGECAATGEC AGEGEATETE TEATTECEGA CAGTGAGAGE GCCETTECTA	READIN	2999292929	Gececerec	GCATGAACTC	ACTCAGAGCA						ပ			
FIG. 6a	<b>ວ</b> ອອອອອວວອອ		TCATGCCCGG	ATTGGCCGCT	CTCCGCTACA GCATGAACTC	GAAAGCATCC ACTCAGAGCA	ACCTGGAGAA GAACGAGCAC	GTGGCAATCC TCAAACCAGC	TGAGGGACGT GGATGAGTCC	TTATAGTGAG AAGTAACCTC	ACCGTATTTA CAGATCCGAA	CCGGCCCAC CTCGAAGGA	GCAGAATGAA ACGGGAAGAT	ATTATGCAGA GTTTCAGAAG	

#### 960 1020 1080 1140 1200 1260 1320 1380 1440 1500 1560 1620 GTGGACGCAT GGGGATAAGT GTGAAGTTTC AGAGAATCCC TACTCTACCC TCTTAGTTGG AGGCGCAAGC TGCTTGCTCA GAAGAGCCAT GACAATGCTC TCACTTACCA GGAACAGTGC CAGCAGCTGT GGGGACCTGG AGCCCGGCCT GCCCTCGATC AGATTGCCAA CTATGTTGAT AAGTTTTACC GCTCCCTGAA CATCCGAATT GCACTTGTCG AGCTAATCAC GGGCAGGTCC TTCCAAGGCA CCACCATTGG CCTGGCCCCC CTCATGGCCA TGTGCTCCGT GTACCAGTCT GGAGGAGTTA GCATGGACCA CTCCGAGAAT GCCATTGGTG TAGCCTCCAC TGTGGCCCAT GAGATTGGCC ACAACTTTGG CATGAGCCAT GATTCTGCAC ACTGCTGTTC TGCCAGTGCA GCCGATGGCG GCTGCATCAT GGCCGCCGCC ACCGGGCACC CTTTCCCCAA AGTGTTCAGT TGGTGTAACA GGAAGGAGCT GGACAGGTAT CTGCAGACAG GTGGCAACGG GTACCTGGAA GACGGTGAAG AATGTGACTG TGGAGAAGAG GAGGAATGTA AGAACCCTTG CTGCAATGCC TCCAACTGCA CTCTGAAGGA AGGGGCAGAG TGTGCCCATG GTICCIGCIG CCACCAGIGC AAGCIGGIGG CICCIGGAAC CCAGIGICGG GAGCAGGIIC GGCAATGTGA CCTCCCCGAG TTCTGCACCG GCAAGTCTCC CCACTGCCCC ACCAACTATT ATCAGATGGA TGGCACCCCC TGCGAGGGTG GCCAGGCCTA CTGCTACAAC GGCATGTGCC GAGGAGGGAT GIGICICICC AACAIGCCGG ACACIAGGAC GCIGIAIGGA GGCCGGAGGI GCTTGGAGGT TCTGGTCCTT FIG. 6b

## FIG. 60

1740	1860 1920	1980	2100	2160	2220	2280	2340	2400	2460	2520	2580
AAGGGCTTGA CAGTGCCAGA	ATCACCTTGA GAGGAGGAAG	GGCCACAACC	Trcccreecr	GGTCCTTTGC	GTGTTGGCAG	CCCTCGGCTC	AGTGGTGGAA	AAGGTGACTA	CCAGACTACC	GCTGGGAGCT	CCCCCAAGCC
AAACTGTGGC TGGSAAGATT	TGACACCACC GGGTCCTGAG	AACCAAGTGT TGAGACGGAA	CTGTCATTGC	CGTCGACAGT	AGCTCTCTTC	ACTGGGCAAA	GGTATCTCAG	GGCCAAGCGA	CCGGCCCCT	GAACAGGGCT	CAGGAGGCCT
ACACCTATGG ATGCCAAGTG	CAGTATCTAT	TGATGACTGG	ACAACAAGAA	ATGGTGGCAG	GGGTGTTTTC	AGAGCCACAA	GTCCCTTCAG	AGACCCCCCA	ACCCCCCTCT	CGGCACATCT	AGGAGTCAGC
GCTGCTGGTG AGTCCCAGGG	GAATCCAACG	CCAGGGCTGG TGCAGGAACA	GTCTGCAACA	ACCCCGGGAG	CC GTGATCGCTG GGGTGTTTTC AGCTCTTTC GTGTTGGCAG	TGCTACAGAC	CAGTTCAGTT GTCCCTTCAG GGTATCTCAG AGTGGTGGAA	TTCAAGTTGC AGACCCCCCA GGGCAAGCGA AAGGTGACTA	AAGCCGTCCC	GCACCATTGT	ATAGAAAGAA
TTTGCTTTGA GAGGGTGAAT GCTGCTGGTG ACACCTATGG AAACTGTGGC AAGGGCTTGA ATGGCCAATA CAGGAAGTGC AGTCCCAGGG ATGCCAAGTG TGGSAAGATT CAGTGCCAGA	GCACCCAGGC CCGCCCCTG GAATCCAACG CAGTATCTAT TGACACCACC ATCACCTTGA ACGGGAGGCG GATCCACTGT CGGGGCACCC ACGTCTACCG GGGTCCTGAG GAGGAGGAAG	GGGAAGGTGA CATGCTGGAC CCAGGGCTGG TGATGACTGG AACCAAGTGT GGCCACAACC ATATTTGCTT CGAGGGGCAG TGCAGGAACA CCTCCTTCTT TGAGACGGAA GGCTGTGGGA	AAAAGTGCAA TGGCCACGGG GTCTGCAACA ACAACAAGAA CTGTCATTGC TTCCCTGGCT	GGTCTCCACC TTTCTGTAAC ACCCCGGGAG ATGGTGGCAG CGTCGACAGT GGTCCTTTGC	Terecerece	TICTGGTGCT ACTGTCTAC TGCTACAGAC AGAGCCACAA ACTGGGCAAA CCCTCGGCTC	TCCCTTTCAA GCTGCGGCAT		ACACCCCTGA ATCCCTCCGG AAGCCGTCCC ACCCCCTCT CCGGCCCCCT CCAGACTACC	TGCGCGTTGA ATCGCCACCT GCACCATTGT CGGCACATCT GAACAGGGCT GCTGGGAGCT	CCCCAGAAGC TGGGGCTCGA ATAGAAAGAA AGGAGTCAGC CAGGAGGCCT CCCCCAAGCC
TTTGCTTTGA ATGGCCAATA	GCACCCAGGC ACGGGAGGCG	GGGAAGGTGA ATATTTGCTT	AAAAGTGCAA	GGTCTCCACC	CCCCTAAGAG TGTGGGTC	TTCTGGTGCT	TCCCTTTCAA	CTGGCCATGC CAACCCAACT	ACACCCCTGA	TGCGCGTTGA	CCCCAGAAGC

3120	GCTACCAGCG	CTAGCCAAGA	CCTCAGCCCA	AGCACGCGA GGAGGTGGG AAAGGTTCTC CCTCAGCCCA CTAGCCAAGA GCTACCAGCG	GGAGGGTGG	AGCACGGCGA
3060	AGGGAAGGGA	CAGCCCAGGG	GTGCAATATA	CATATCCAAT GECTCCTAAG TGTTTGTCCT GTGCAATATA CAGCCCAGGG AGGGAAGGGA	GGCTCCTAAG	CATATCCAAT
3000	CAGGGGGACT	ACCCGAGTGC CTCCTGCTTC CTCCGAGGCC CAGGGGGACT	CTCCTGCTTC		TCCCCAGAAT CTCCCTTCTT	TCCCCAGÁAT
2940	ACGTCCCTCA	GCCTTCTGGA	AGCTTTGGCA	AGAAGAAAGA CGCCTTCTCA CCCATCCTGA AGCTTTGGCA GCCTTCTGGA ACGTCCCTCA	CGCCTTCTCA	AGAAGAAGA
2880	TGGAAGGTCC	CCGGATGCCA	ATGGAAGACT	TCTAGAAGTG TCGAGAAGTT TCTTGTTCCG ATGGAAGACT CCGGATGCCA TGGAAGGTCC	TCGAGAAGTT	TCTAGAAGTG
			•	:		
2820	AGCTCCAAGA	AGCAATAATT	AGAGGGTTGG	CTGTTCCAAA GCTACCCGAG TACCGATCAC AGAGGGTTGG AGCAATAATT AGCTCCAAGA	GCTACCCGAG	CTGTTCCAAA
•	ME	READING FRAME	REA			
2760	CCAGCAGTGC	CCCCAGGCCT	GTCCTCAGCC	GCACCTCCCT GCTTCAGCCC CCTACTTCTG GTCCTCAGCC CCCCAGGCCT CCAGCAGTGC	GCTTCAGCCC	GCACCTCCCT
2700	AGGTCAGGAG	GACCGGTCCC	CAGGCCAAAG	CTCAGAAGGC ACTCCCAGCC AATCCGGTGC CAGGCCAAAG GACCGGTCCC AGGTCAGGAG	ACTCCCAGCC	CTCAGAAGGC
2640	CCTCGACCAC	CTTCTCCAGG	TGTCCCAGGA	GACCCATGCC CCCTGCACCT AACTGCCTAC TGTCCCAGGA CTTCTCCAGG CCTCGACCAC	CCCTGCACCT	GACCCATGCC
						FIG. 6d

FIG. 6e						
ATGCTCAGGG	AAGGCTTGAG	ATGCTCAGGG AAGGCTTGAG CTGGGGTCCT CCTCTGCGGA GCTTGCAGAA GGTACCCATC	CCTCTGCGGA	GCTTGGAGAA	GGTACCCATC	318(
CTGGTCCTAT	GCTGGCAGGA	CIGGICCIAI GCIGGCAGGA ACACACGCGA GIGICACIGA IIGGCCICCI ICIGGGAICC	GTGTCACTGA	TIGGCCTCCT	TCTGGGÄTCC	324(
CAGGCTGCTG	AGGAAGCTAC	CAGGCTGCTG AGGAAGCTAC TGCTACATCC CTACCCCAAG GGGCTTGGTC AAGGTGCCTG	CTACCCCAAG	GGCCTTGGTC	AAGGTGCCTG	330(
TYCCTGGCTC	TCTGGCTGCA	TYCCTGGCTC TCTGGCTGCA TGTAATAAGC CATGCTCCCC TCCCCTGCCT TTCTTCACAT	CATGCTCCCC	TCCCCTGCCT	TTCTTCACAT	336(
TCCCACTCCC	ATATTTACAC	TCCCACTCCC ATATTTACAC GGGTCACTCT GACTCAGACA GGTACTATTT	GACTCAGACA	GGTACTATTT	GTAAGTAGCA	342(
TAGACAGCAG	GGGGGTGGGG	TAGACAGCAG GGGGGGGGG TGGTCAACCT GTGTCCCCTC TGAGCCGTTA TGCCAAAGGT	GTGTCCCCTC	TGAGCCGTTA	TGCCAAAGGT	348(
CACTAAGGAC	ATTTAGAATC	CACTAAGGAC ATTTAGAATC CCCATCCATC CATCCATCCA TCCATCCATC CATCCATTCA	CATCCATCCA	TCCATCCATC	CATCCATTCA	354(
TCCATCCCCA	GTGTTCCATG	TCCATCCCCA GIGITCCATG IGICACCITC ICCITITCCA GCATCCCIAT CCTAIGGIGC	TCCTTTTCCA	GCATCCCTAT	CCTATGGTGC	3600
TTTGGTGGTG	AACTATGGCA	TTTGGTGGTG AACTATGGCA GTCCTGACTT GCTGATGACC ATATGCTGGT GACCTACAAA	GCTGATGACC	ATATGCTGGT	GACCTACAAA	3660
TCGGGATCCT	GCCATATGGG	TCGGGATCCT GCCATATGGG GTCGCCACTG GACTTTCTGC ACTGGTTCTC AAGAGCGTTG	GACTTTCTGC	ACTGGTTCTC	AAGAGCGTTG	372(
AGCCGAGTGG	GCGTGTATGT	AGCCGAGTGG GCGTGTATGT TTGTGTGT GTGTGTGT GTGTGTGTGT GTGTGTGTGT	GTGTGTGTGT	GTGTGTGTGT	GTGTGTGT	378(
GTGTGTGTGT	GTGTGTGTGT	GTGTGTGTGT GTGTGTGT GTGAAAGAGA CAGAGGCAAT GAGAGAGACA GACATGCAGG	CAGAGGCAAT	GAGAGAGACA	GACATGCAGG	384(
CAGGCCGACA	GCTCTGCATG	CAGGCCGACA GCTCTGCATG TACTTGTGTT TTACGCCCTC AAGCAGTATA AGGGACCTCC	TTACGGCCTC	AAGCAGTATA	AGGGACCTCC	390(
TCCTTATTTC	TGACTCATAT	TCCTTATITC TGACTCATAT CTAAGTAAGG TTCCCCAGGA CMAGCCACAG CTGTACTGAG	TTCCCCAGGA	CMAGCCACAG	CTGTACTGAG	396(
GGGGCTGAC	ATGTTTGGCA	GGGGGCTGAC ATGTTTGGCA TCCTGGCTAT AGTATTGTAT ACACAGGGCC ACCAGCCCCG	AGTATTGTAT	ACACAGGGCC	ACCAGCCCCG	402(

# FIG. 6f

A 4080	A 4140	G 4200	T 4260	T 4320	T 4380	T 4440	G 4500	T 4560	A 4620	T 4680	G 4740	T 4800	T 4860	T 4920
TCAACAGCC	TTATCACGC	TGCTGAGAA	GAACAGTTC	AGGTGCTAA	TGTTAGCTC	GAGGCCTCC	GACAGACGT	CCGTGCTCT	CTCTTTGTG TGCTTTACCA	CCTCCACTT	ATCACTTCT	TGTGANGTG	CACACAACT	CCCTTCCCA
CAGATCGATG	AACCAGAGTT	TCATGGAAAA	CCGCTATTTG	ACCCCAGAA	AGCAAGCAGT	TTTACCACCT	TGGGAGCTAT	TTCAGGGTGA		AGTTAGAAAC	CTCTTTGGTA	GACATTATTC	AGCTGCTCTA	AGGCAGGAAT
TGACTCTGAA	AACTCTATTC	GAGGTGATCG	GCTTGGTGCT	GCCATCTTCC	CCGCCCCAAC	TGACTTCCCC	TGTAGACATG	CTTTCTCCTT	ATATCAGTTT	CTCACCACAA	CATTTCCTTG	TCCATCCCAT	ANATGTCGGC	AGGCTGTCAC
CCCTAGTGGT CAGCTCTGAG GGGGGACTGG TGACTCTGAA CAGATCGATG TCAACAGCCA	TGGTGAACCA GATCTGGGCA GGGTTCCCCA AACTCTATTC AACCAGAGTT TTATCACGCA	TIGCIGCCCC	ACCTTCTCT	TACACATITG CIGGGCCIGG CCTCTGAGAG GCCAICTICC ACCCCCAGAA AGGIGCTAAI	CT AGGGGCCTCC CCGCCCCAAC AGCAAGCAGT TGTTAGCTCT	GGCAAGCGTT	AA GCTTTGGGAT TGTAGACATG TGGGAGCTAT GACAGACGTG	AGGAAAGCAC	TCCATGTCCT	CACCCCGATT	CTGTCCCTTG AACCATATCA GAAAAAGACC CATTTCCTTG CTCTTTGGTA ATCACTTCTG	TTTTTTCTTC TTCATTACTG TGCTACCACC TCCATCCCAT GACATTATTC TGTGANGTGT	AAGAGGACGG TGTTTNTTA NTCTTGGGAG ANATGTCGGC AGCTGCTCTA CACACATT	GAGGCCAGCT
CAGCTCTGAG	GATCTGGGCA	TCTCTCCTGG	GATGGGGTGG	creeeccree	AGGCCTCTCT	CAGAGGAAGA	CCCAGAGTAA	GAAAGATCTC	GAGGCCTCAG	GACTACAGGC	AACCATATCA	TTCATTACTG	TGTTTTTTA	TTTGTCTCCA
CCCTAGTGGT	TGGTGAACCA	NCICATCGGG ICICICCIGG TIGCIGCCCC GAGGIGAICG ICAIGGAAAA IGCIGAGAAG	GTGGGAATGG GATGGGGTGG ACCTTCTCTT GCTTGGTGCT CCGCTATTTG GAACAGTTCT	TACACATTTG	GGCACTGCAG AGGCCTCT	TGGAACCCTC CAGAGGAAGA GGCAAGCGTT TGACTTCCCC TTTACCACCT GAGGCCTCCT	TATATCTCTT CCCAGAGT	GCCTGGGGTA GAAAGATCTC AGGAAAGCAC CTTTCTCCTT TTCAGGGTGA CCGTGCTCTT	CACACTCTCT GAGGCCTCAG TCCATGTCCT ATATCAGTTT	AGTGGCCGGT GACTACAGGC CACCCCGATT CTCACCACAA AGTTAGAAAC CCTCCACTTT	CTGTCCCTTG	TTTTTTCTTC	<b>AAGAGGACGG</b>	CACTCAAGGC TTTGTCTCCA GAGGCCAGCT AGGCTGTCAC AGGCAGGAAT CCCTTCCCAT

110.0g						
CTGCTTTGTG	AAGGGTCCCA	CTGCTTTGTG AAGGGTCCCA TACAGGTGTA TCTAGACTTC AAGGACAGGG TTTGTCTCAC	TCTAGACTTC	AAGGACAGGG	TTTGTCTCAC	4980
AGGATTGTCA	CTTAGGAGAT	AGGATTGTCA CTTAGGAGAT GAAAGAATAT TACCACATGA GGAGGAGGGG CAGTTGCAAC	TACCACATGA	GGAGGAGGGG	CAGTTGCAAC	5040
AGAACACTTT	GGTCTTCCTA	AGAACACTTT GGTCTTCCTA CACCAAGTCT GTGAGGGCAT CCAAGACTGA ATGAAAGCGC	GTGAGGGCAT	CCAAGACTGA	ATGAAAGCGC	5100
TTTTCTTATG	CATACAATGT	TITICITATG CATACAATGI GAGCAAGAAC AAGAACTGII TAAGGCACCI CIGIICCCAG	AAGAACTGTT	TAAGGCACCT	CTGTTCCCAG	5160
CCACTGAAGA	GAGACGTCAG	CCACTGAAGA GAGACGTCAG AAGATGTTAG AATAGGTCAA AACCAAGGCT CTGGTGGACT	AATAGGTCAA	AACCAAGGCT	CTGGTGGACT	5220
GAGGGAAGGT	TTGTAGCTGC	GAGGGAAGGT TTGTAGCTGC GTTTAGTGGT ATACATCTTT AGTCCCAGCA TAGGCAGGTG	ATACATCTTT	AGTCCCAGCA	TAGGCAGGTG	5280
AATCTCGAGT	TTGAAGCTAG	AATCTCGAGT TTGAAGCTAG CCTGGTCTAA AAAGGAAGTT CCAAGACTGC CAGGGCCACA	AAAGGAAGTT	CCAAGACTGC	CAGGCCACA	5340
CAGAGGAAAA	AAAAAACCC	CAGAGGAAAA AAAAAACCC TCTAGAAAAA CAAAAATGAA GACAGGTTCT CATGTATCGT	CAAAAATGAA	GACAGGTTCT	CATGTATCGT	5400
AGATTGGCCT TTAAGTCA	TTAAGTCACT	CT TTACCAAGGA TGATCTTTGA ACTCCTGAGT ACAGACTGCG	TGATCTTTGA	ACTCCTGAGT	ACAGACTGCG	5460
GGTGTGTGCT ACCATGCT	ACCATGCTTT	ATGTGGCCCT	ATGTGGCCCT GGGTTCAAAC ACAGCCCTTC ATATGTATAT	ACAGCCCTTC	ATATGTATAT	5520
AGCCAAACAC	TCTACAACTG	AGCCAAACAC TCTACAACTG AGCTACATCC TCCAGCCTAG GCTGTAAATG TTTTTTGGAG	TCCAGCCTAG	GCTGTAAATG	TTTTTGGAG	5580
CTAGATTAGC	TGCCTGCCAA	CTAGATTAGC TGCCTGCCAA CCTTAGAACT GCAAAGCCAT TCCTGACCTG TAAACCTCAG	GCAAAGCCAT	TCCTGACCTG	TAAACCTCAG	5640
CTCTCCATCT	CTATAAGAGG	CTCTCCATCT CTATAAGAGG TATAGCCTGG GCTAATACCG TCCAAGTTAC AACTCCTTGC	GCTAATACCG	TCCAAGTTAC	AACTCCTTGC	2100
TTGCTTTCTG	TTCCTTCTAG	TTGCTTTCTG TTCCTTCTAG CCTTGGTGAC TTCCACCAGG AAGAGAATAC CCCCTCTCTA	TTCCACCAGG	AAGAGAATAC	CCCCTCTCTA	2160
CCCCTGCTCC AAGACACT	AAGACACTGT	GT AGATGCTAGT GTCGGAGTGT TCTCTGTAAC GCGACAGTTC	GTCGGAGTGT	TCTCTGTAAC	GCGACAGTTC	5820

#### 5940 0909 6180 6240 6352 0009 6120 6300 TITITIGITG IIGITGITAT IITITICICC AGTTATTTTG TCAAATGCAT GTAATGAACA GACCCGAAGG AATCCTCCAC ACACAAGCCA GGGAACACCA ACTGGAAAGG TACCCCGTCC CAGGGAAGCC TGCTAGGGAG AGGTTCTGTA GAATCCGAGC CTAGCACCCC AAAGTCATGC ACCCAGTATC CTCTTGTATG ACTGTATATG ATTTGGGAGA TTGTTCACAG ACAATATTG TACATCTATG CTTCTGTTGC AATAGCCCCC CTGCAACACT GCAATAATCC TTCAGTGTCT CCCCTGGGCT CAATTCACTT CCTTATTTGA CAAAGTGGAG GTGAGACTTG TATTCTTAAA ATTGGAGGCT ITGTTTGTAT TTAATTAAAA CAAATTGTCA TGAGGAAAAA AAAAAAAAA AA ICTATGICIG GGAICCAGGG CAAAIGIGAA IIICCIIIIG GAAGTAGICC ICCCCICICA IGICCICCIA IIGAIIGIII CAAAATACTT GAATGGGCCA TGGTGCCTTG

FIG. 6h

TIGITGIAGA CAAGGAAAGG TACGACATGA TGGGACGGAA CCAGACTGCT GTGAGAAG

CTCAGCTGCT GCGCAGAAGA AGAGCTGTTC TACCACAGAC CCGCTATGTG GAGCTGTTCA

FIG. 7a			•		٠	
GTTGCAAGGA TGACCGAAGN NCGGAGGCGG CGGCCGCGCG TTGAGCGGAA CCTGCCGAAG	SAAGN	NCGGAGGCGG	9292920992	TTGAGCGGAA	CCTGCCGAAG	
	REA	READING FRAME	ME			
 CCTCGCTAT GGGGCGGCG GCGCTCTCGC CCCTTGCCTC TCTGCGACTA AGGTGGCTGC	29292	cccrcrcc	CCTTGCCTC	TCTGCGACTA	AGGTGGCTGC	
TEGESTETES CTTSCTESSC CCASTCCTCS AGGCCGGCC ACCAGACTTS GAACAGACTG	reeec	CCAGTCCTCG	AGGCCGGGCG	ACCAGACTTG	GAACAGACTG	
TCCATCITIC ITCTIATGAA ATTAITACTC CITGGAGAIT AACTAGAGAA AGAAGGGAAG	ATGAA	ATTATTACTC	CTTGGAGATT	AACTAGAGAA	AGAAGGGAAG	
CICTGGGGCC CAGITCACAG CAGAICTCTI ACGICAICCA GGCCCAAGGA AAACAGCAIA	CACAG	CAGATCTCTT	ACGTCATCCA	GGCCCAAGGA	AAACAGCATA	
TTATTCACTT GGAAAGAAAC ACAGACCTTT TACCTAATGA TTTTGTAGTT TACACCTAGG	SAAAC	ACAGACCTTT	TACCTAATGA	TTTTGTAGTT	TACACCTAGG	
ACAAGGAAGG CTCCCTACTC TCTGACCATC CCAACGTACA GAGCCATTGT CACTATCGAG	FACTC	TCTGACCATC	CCAACGTACA	GAGCCATTGT	CACTATCGAG	
GCTATGTGGA GGGAGTGCAG AATTCCGCGG TTGCTGTGAG CGCCTGCTTT GGACTCAGAG	rgcag	AATTCCGCGG	TTGCTGTGAG	CGCCTGCTTT	GGACTCAGAG	
GCTTGCTGCA TTTGGAGAAT GCCAGTTTTG GAATTGAACC TCTGCACAAC AGCTCACACT	AGAAT	GCCAGTTTG	GAATTGAACC	TCTGCACAAC	AGCTCACACT	
TIGAGCACAT ATTITACCCC AIGGAIGGCA ICCACCAGGA GCCICIGAGA IGIGGAGICI	3000	ATGGATGGCA	TCCACCAGGA	GCCTCTGAGA	TGTGGAGTCT	
CTAACAGGGA CACAGAGAAG GAAGGCACAC AGGGGGATGA GGAGGAGCAT CCGAGTGTCA	AGAAG	GAAGGCACAC	AGGGGGATGA	GGAGGAGCAT	CCGAGTGTCA	

## FIG. 7b

840	006	096	1020	1080	1140	1200	. 1260	1320	1380	1440	1500	1560	1620	1680
ATTCGAATTG	GGAGGAGCTG	CGTCGGAGAC	ATGGCGTTTG	CANATCACTG	ATGAATCATG	GGAGCATCCG	TTGAATAAGG	cccrccrere	GAGTGTGAGG	TGTGCATATG	GGGAAGACCA	CCAGATGTCT	GGCATGTGCC	GCCCCAAGAG
CATGTTAAAC	CAATATAATT	CCTTATAACT	AACTGCAGGA	Tetettteeg	TAACCTTGGA	CATGAATTCA	GAAGTTAACG	CTACAGCGCG	CACAGCGAAG	ATTTGCTGAG	CATGTGCAGA	GTTCTGCCCG	CTGCTACAAT	GGCTAAGGCT
GCATGTACAT	GAAATCCTAT	GGGAAAAGTT	GCTTTGGTGG	GTGGGATCAA	AATTGGGGCA	AGAGCTGTAT	AGGACTTTGA	CTGACGAAGC	GTGACTGCGG	AGCTCAAGTC	CAGGAGGCTC	GTTCCTCTCA	GCAAAGCCTA	TTGGTTCAAA
AGATGATTCG CTTAGCAAAC TACCTGGATA GCATGTACAT CATGTTAAAC ATTCGAATTG	TGCTGGTTGG ACTAGAAATT TGGACAGACA GAAATCCTAT CAATATAATT GGAGGAGCTG	GTTCAGTGGC GGGAAAGTT CCTTATAACT CGTCGGAGAC	ACGACAGTGC ACAGTTGGTT TTGAAGAAAG GCTTTGGTGG AACTGCAGGA ATGGCGTTTG	AGCCACGCAG	ITCCATT GTTGCTCATG AATTGGGGCA TAACCTTGGA ATGAATCATG	TGTGGAGCAA	TGCAGTGCGG	ATCCCGAAGC	GGAGAGGAGT	AGCACTTGTA	CAGTTCCTTC	TACTGCAACG	TGCCAGAACA	CAGGTCATCT
CTTAGCAAAC	ACTAGAAATT	GAGATGTGCT GGGCAACTTT	ACAGTTGGTT	ATGTTCAAGG	TGCATCCATT	AGAGTGTTTC	CTTTAGCAGT	CCTGCTTAAC	GGTGGACCCT	CTGTGAAGGA	TAAAGATTGC	TGTTCCTGAG	TGGATATCCT	CGCGCAGTGT
AGATGATTCG	TGCTGGTTGG	GAGATGTGCT	ACGACAGTGC	TAGGAACAGT ATGTTCAAGG AGCCACGCAG GTGGGATCAA TGTGTTTGGG CAAATCACTG	TGGAGACATT TGCA	ATGATGGGAG AGAGTGTTTC TGTGGAGCAA AGAGCTGTAT CATGAATTCA GGAGCATCCG	GGTCCAGAAA CTTTAGCAGT TGCAGTGCGG AGGACTTTGA GAAGTTAACG TTGAATAAGG	GAGGAAGCIG CCIGCITAAC ATCCCGAAGC CTGACGAAGC CTACAGCGCG CCCTCCTGIG	GTAATAAGCT GGTGGACCCT GGAGAGGAGT GTGACTGCGG CACAGCGAAG GAGTGTGAGG	TGGACCCATG CTGTGAAGGA AGCACTTGTA AGCTCAAGTC ATTTGCTGAG TGTGCATATG	GCGACTGTTG TAAAGATTGC CAGTTCCTTC CAGGAGGCTC CATGTGCAGA GGGAAGACCA	GTGAGTGTGA TGTTCCTGAG TACTGCAACG GTTCCTCTCA GTTCTGCCCG CCAGATGTCT	TCATTCAGAA TGGATATCCT TGCCAGAACA GCAAAGCCTA CTGCTACAAT GGCATGTGCC	AATATTATGA CGCGCAGTGT CAGGTCATCT TTGGTTCAAA GGCTAAGGCT GCCCCAAGAG

#### 2220 2040 1980 2280 1800 1860 1920 2100 2160 2340 2400 2460 2520 TCTCAGGGAA ACTTGATTCC GGCTCGGCCC GCTCCTGCAC TGAAGTCAAT TCTAAAGGTG ACAGATTTGG CAACTGTGGT TTCTCCGGCA GAAGTGTGCC ACTGGGAACG CGCTGTGTGG AAAGCTTCAA TGCGAGAATG GGGTGTGGAT TTCCAGCTTG GTTCCGACGT TCCAGACCCA GGGATGGTGA TTATGACTGT GACATTCAGG GAAAATGTCA TGGCCATGGG GTATGTAACA CGTGGACAGC GGGCCGACGT ATAATGCAAA GAGCACAGCA CTGAGGGACG CITCITCITC CTAAICGICC CCCIIGITGC GGCIGCCAIT IICCICITIA TGAACTACGG AAAACCTTCA GGAAGAAGAG ATCACAAATG TCAGATGGCA AAACGTCTCT AGACAGCCAG GAGATCCTAG TATCTCCAGA CCACCAGGGG CTCCAGACCA CCAGGGGCC CAGGTGTCTC CAGACCACCA GGGGGCCCAG ACCACCAGGG GGCCCAGGTG TCTCCAGACC GCCACCTGGG CATGGAAACA ACCAACCTAC GCGCCAAGC AGCCTGCGCA GTTCCCGTCA AGGCCACCTC TCAGACACCC AGTCGAGGCA TIGICACIGI GAAGAIGGCI GGGCICCCCC ACACTGIGAC ACCAAAGGAI CAAATGTGAT GCTGGCAAGA TTTGCAGGAA TTTTCAGTGT GTAAATGCTT GCCGGTGTTT GGAATAGTAC CAGCTATCAT CACCACAACC GAAAATATCT ATTGCTTCAT ATGAAGGCAC CTGTCCTGAA GCAATAAGAA ATGGAGGAAG GAAATCAAGC GTGTCTCCAG CCAAATGCTG **TCAAGAGAGA** GCCCAAATGT GATTCCCAGT GCCTTCTGGT GTGAGTACAA **TACAGGACAT** FIG. 7c

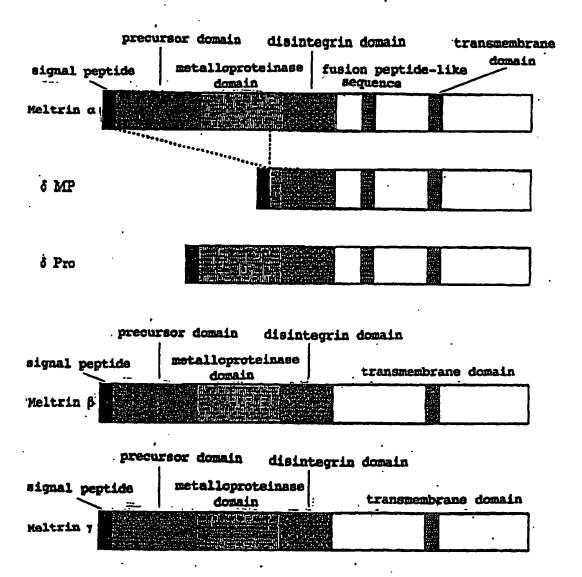
# FIG. 7d READING FRAME ——

C 2640	T 2700	G 2760	G 2820	T 2880	T 2940	3000	T 3060	c 3120	A 3180	A 3240	T 3300	T 3360
TTTAATGT	TGAAAAGCC	AACACAGGA	AAAGTGAAO	GACGTGGAT	TGTAAGCTT	ATAGTTTCT	TCTGCATTT	ACAAAGAAG	TACTTGGAA	TTCAACTAT	TCAAATATC	TATGCATCA'
AATCTTATTT	GATGITITCI	AAGCTTTATT	CGGGGGGTGT	TAATGCACTT	ACTGTATTAG	GTTACCATTA	GCCAATAACG	CTCTTAATGC	GATCATGATA	TTWATGATCT	GCTTCCATTT	CAATTTTCTT
AGAATATTAG	TTTTTTCCT	ACCACAAAAC	GCAGGCGCTC	TTTGTGGATT	TCTCARATTA	TTTGACCCCA	AAGACTAATT	GGACTTATGT	ACAATTACTA	GCTTCCATTT	TCAGTTATGG	ATTCAACTCT
ACCTGATAGT	TITGITGIT	AACACAAACC	ACAGGAATGT	TCTCTGGCCA	CTGTAATGAT	ATCCTGACTT	TTAACTTAGG	CCATTTATAT	CACAAGAACC	TCAGTTATTG	TTTAACACAA	ACAGGAACTG
CICCITIAIA TAGCICCCIC ACCIGATAGI AGAATATIAG AATCITATIT ITTAAAIGIC	TTCAGGGAAC TGAGCAAATG TTTGTTGTTT TTTTTTTCCT GATGTTTTCT TGAAAAGCCT	TICICITICA ACCATGAATG AACACAAACC ACCACAAAC AAGCITITAIT AACACAGGAG	CCTAGTGGGG ATTGCGAAAC ACAGGAATGT GCAGGCGCTC CGGGGGGTGT AAAGTGAACG	TITCCATOGI TAGAATGITI TCTCTGGCCA TITGTGGAIT TAATGCACIT GACGTGGAIT	AAGTTATICT GAGCATGTTA CTGTAATGAT TCTCAAATTA ACTGTATTAG TGTAAGCTTT	GICACTATGC GCIAAACGIA ATCCTGACIT TITGACCCCA GITACCATIA ATAGITICTG	GTTGACCATT TGAACATGTA TTAACTTAGG AAGACTAATT GCCAATAAGG TCTGCATTTT	CATCTTGCAT GGATTAACAG CCATTTATAT GGACTTATGT CTCTTAATGC ACAAAGAAGC	AGATATCTCG AAGGAGCTTA CACAAGAACC ACAATTACTA GATCATGATA TACTTGGAAA	GIGIGAAATA IGGIGIGIAC ICAGITAIIG GCIICCAIII IIWAIGAICI IICAACIAIA	ACAATTATGA TAGAAATCGA TTTAACACAA TCAGTTATGG GCTTCCATTT TCAAATATCT	TITCAACTGT AATGACTATG ACAGGAACTG ATTCAACTCT CAATTITCTT TATGCATCAT
CTCCTTTATA	TTCAGGGAAC	TTCTCTTCCA	CCTAGTGGG	TTTCCATCGT	AAGTTATTCT	GTCACTATGC	GTTGACCATT	CATCTTGCAT	AGATATCTCG	GTGTGAAATA	ACAATTATGA	TTTCAACTGT

#### 3600 3660 3720 3780 3840 3900 3480 3540 3931 TTGTGTATAT ATACATATA AAATAAAAAC ATTTACAACA AATAAAATAC TGAGGTGTTA CCAAACCACT TGAGAATTCA TGAGCACTTT AACTCTAAAC TCTGAATTTC GIGAAGICCI CIAGAAIGII TACAITIACI AAGGIGIGCI GGGICCIGIC TCTTTTGACT AATATTTTCG TAAACATTAG GCTGGAGAAA GGAAGGAAGC AGTGGTTTCC TTAGATAACT ACAGAATTAT ACTGGTCTCT GGGATTACTC TCTCAGCTGT ATTAAAATGA TGAAAGGAAT GATATTGACA CTAAAATTTT AAACATTTAA ATTTTTTCAT AAAGAAGTIT AATAATAGGT ATATTAACTG AATTICATIA GITTITTAAA GGTAAAGCAT TGCAGCAGTG TTGTTTTGTT TGAAGTGCAC ACTCTATGGT ACGAGGTGTT TAGTATACCC AAGCAGATAG GTGTCGATCG AACAGGAGCA GGGAGAATAC TTCCAACAGT TTGAAATTCT AAAAAAAA AAAAAAAA ATTTGTACTT AATCTTTCAT AAAGCTTGAT ATAATATTGT

FIG. 7e

FIG. 8



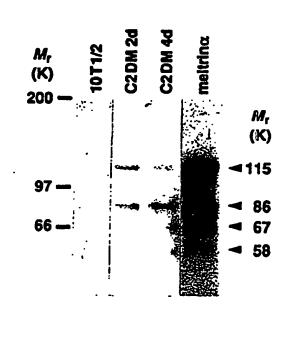


FIG. 9

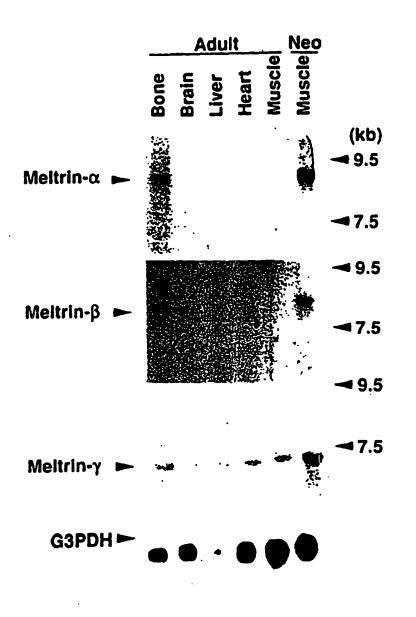


FIG. 10

FIG. 11a

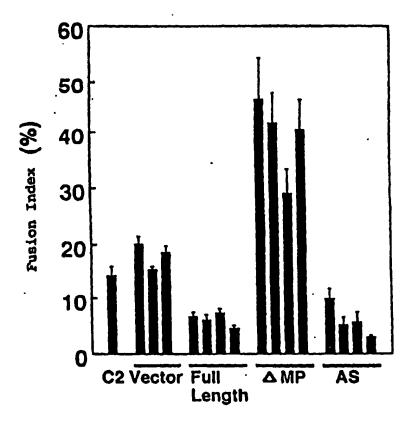
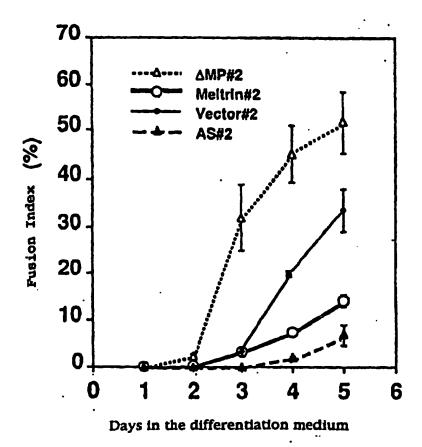


FIG. 11b



	) <u>8</u>			4	8	
250 260 270 280 290 300 GGTGAACCTTATGGCAACTGTGGGAAAGTCTCGAAGAGTTCCTTTGCCAAATGCGAGATG	190 200 210 220 230 240 CTCTGGGGACCAGGTGCTAAACCTGCCCTGGGATCTGCTTTGAGAGGTCAATTCTGCA L W G P G A K P A P G I C F E R V N S A	GATGTGGACGGCTACTGCTANAATGGCATCTGCCAGACTGCAGCAGCAGTGTGTCACG D V D G Y C X N G I C Q T H E Q Q C V T	180	ACAGGGCCAGCCTCACTGCCCAGCCAACGTGTACCTGCACGATGGGCACTCATGTCAG T G A S P H C P A N V Y L H D G H S C Q	AAGCCTGCAGGAACAGGACTCCAGCAACTCCTGTGACCTCCCAGAGTTCTGC	9.
· CGA	TTC	TGT V		ATG C	GTT F	
ATG	CAA	ဥ္ ၁		CTC S	AGA E	
290 CAAAT	230 AGT( V	y ပွဲ	170	GCA(	200	50
TGC	GAG	Q Q		ງີດີ ດ	CCT	
CTT	TGA E	E E	•	CGA'	TGA D	
280 GTTC	220 GCTT	TCA H	160	TGCA(	CTG	40
GAG.	CTG	GAC	Ā	CCI	CTC S	•
GAA	200 210 220 230 24 GGTGCTAAACCTGCCTTGGGATCTGCGC G A K P A P G I C F E R V N S A	₹ 22 27	!	GTA. Y	ACAGCGTGCAGGACTCCTGTGACCTCCCAGAGTTCTGT A C R D S S N S C D L P E F C	
CTC	0 IGG	ည် သ	0	<b>V</b>	CAG S	0
270 AGTCT	210 CCCTG( P G	CAT	150	CAAC	S S	30
CAA	TGC A	76G G		AGC	GGA	
766	ACC' P	NAN		ည်	CAG R	
260 CTG	200 TAA/ K	CTA ×	140	S	GTG C	20
CAA	TGC'	S S S		TCA H	AGC A	
<b>76</b> 6	AGG' G	CTÀ ≺		ည္သ	AAC, T	
250 CTTA'	190 CTCTGGGGACCA L W G P	GATGTGGACGGCTACTGCTANAATGGCATCTGCCAGACTGTGTGTCTCA( D V D G Y C X N G I C Q T H E Q Q C V T	130	ACAGGGCCAGCCTCACTGCCAGCCAGGTGTACCTGCACGATGGGCACTCATGTCA T G A S P H C P A N V Y L H D G H S C Q	AGG, G	01
ACC	1. GGG, G	GGA D	=	095 A	TGC, A	}   
TGA	CTG ₩	TGT. V		AGG.	AAGCCTGCAGG K P A G	)
33	CI	GA.		AC. T	AA!	7 1 1 1 1 1

FIG. 12b 310 320 AGAGATGCTAAATGCGGCAAG R D A K C G K 10

20	40	09	80	•
20 30 40 50 60 ATCATGAATTCAGGAGCATCGGGTTCCAGAACTTTAGCAGTTGCAGT I M N S G A S G S R N F S S C S	70 80 90 100 110 120 GCAGAGGACTTTGAGATTAAATAAGGAGGAAACTGCCTTCTTAATATTCCA A E D F E K L T L N K G G N C L L N I P	130 140 150 160 170 180 AAGCCTGATGAAGCCTATAGTGCTCGCTGGGGAA K P D E A Y S A P S C G N K L V D A G E	190 200 210 220 230 240 GAGTGTGTGTACTCCAAAGGAATGTGAATTGGACCCTTGCTGCGAAGGAAG	250 260 270 280 290 300 TGTAAGCTTAAATCATTTGCTGAGTGTGTGTGTTGTAAAGACTGTCGGTTC
77.G C	IAT		AAG S	250 260 270 280 290 30 TGTAAGCTTAAATCATTTGCTGAGTGTGCATATGGTGACTGTTGTAAAGACTGTGGTT
AGT S	Z. W	A A	999	TG
50 AGC S	110 TCTI	170 GGAC D	230 CGAA E	290 AGA(
TTT F	CTI	1 GTG V	2 760 0	AAA
AAC	16C	TTG L	76C C	TGT
40 CAGA R	AAC N	O AAG K	o CCT	o TGT
TCC,	100 GGAAA G N	160 TAATAAGTT N K L	220 GACCC D P	280 GACT(
GGT	3GA(	140 150 160 170 18 GCCTATAGTGCTCCCTCGTGGTAATAAGTTGGTGGACGCTGGGGAAAAAAAA	200 210 220 230 24 GGTACTCCAAAGGAATGTGAATTGGACCCTTGCTGCGAAGGAAG	36T
rc6(	YAA	rgr C	3AAY.	rat(
30 30 30 30 30	90 VAT/	150 2TCC7 S (	210 ITGT( C I	270 GCAT
GA(	TTA.		3AA.	, je
CAC:		SCT6	AGG	. AG
20 AATI N S	80 TTAA L 1	140 TAGTC S A	CAA	260 TGCT(
TGA	AGT	14 ATA	200 CTCC	26 TTC
TCA	AGA		GTA	CAT
GCA	TTG	AAG	GTG G	AAT
10 GCTG C	70 ACTI F	130 3ATGA	190 ACTG	250 TTA
AGA S	1GG	CTG D	3TG D	AGC
10 20 30 40 50 GCAAAGAGCTGCATCATGAATTCAGGAGCATCGGGTTCCAGAAACTTTAGCAGTTGCAG A K S C I M N S G A S G S R N F S S C S	70 80 90 100 110 12 GCAGAGGACTTTGAGGTTAACTTTAATAAGGAGGAAACTGCCTTCTTAATATTCC A E D F E K L T L N K G G N C L L N I P	AAGCCTGATGAA( K P D E /	GT(	ïTA!
GC A	QC A	¥ ¥	යි ස	75

120	140	160	180	200
STO 340 350 360 360 360 360 360 360 CTTCCAGGAGGTACTTTATGCCGAGGAAAACCAGTGAGTG	370 380 390 400 410 420 AATGGTTCTCCAGTTCTCTTTTTTTTTCAGAATGGATATCCTTGCCAG N G S S Q F C Q P D V F I Q N G Y P C Q	440 450 460 470 480 ITTGCTACAACGCCATGTGCCAGTATTATGATGTCAAGTC C Y N G M C Q Y Y D A Q C Q V	490 500 510 520 530 540 ATCTTTGGCTCAAAGGCTGCCCCCAAAGATTGTTTCATTGAAGTGAATTCTAAA I F G S K A K A A P K D C F I E V N S K	550 560 570 580 590 600 GGTGACAGATTTGGCAATTGTGGTTTCTCTGGCAATGAATACAAGAAGTGTGCCACTGGG G D R F G N C G F S G N E Y K K C A T G
r r	55	CAA.	ي کي ر	CT
AG.	55	GTC	ATT	
CAG	410 ATATC Y F	0 Y	0 TGA	o GTG
350 TTCC/ P	41 GAT Y	470 CTCAA	530 AAGT( V	590 AGTG1 C
ATG V	ATG G	A TG	7. E	\GA K
TG/	GA/ N	TG/	CAJ	CA/
340 AGTG	400 TTCA	460 'ATTA	520 GTTT	580 AATA
3 TGA E	4 TAT I	4 GTA	5 ITG C	5 TGA E
SAG	T.T. F	SCA CA	AGA' D	CAA
LAC(	CGT;	STG(	CAA K	
330 AAA/ K	390 GAT	450 ATG1 M	510 CCC, P	570 TCTG S G
	380 390 400 410 42 GTTCTGTCAGCCAGATGTTTTTTTTCAGAATGGATATCCTTGCCA F C Q P D V F I Q N G Y P C Q	) ) )	500 510 520 530 54 GCCAAGGCTGCCCCCAAAGATTGTTTCATTGAAGTGAATTCTAA A K A A P K D C F I E V N S K	560 570 580 590 60 AATTGTGGTTTCTCGGCAATGAATACAAGAGTGTGCCACTGG N C G F S G N E Y K K C A T G
CGA R	CAG Q	AAC	3CT A	3GT G
320 ATGC	380 CTGT C (	440 CTAC	500 CAAG	560 TTGT
35.	33 33 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5	4. 25.	200	5( VAT;
CTI	AG1	'ATI	AA .	759.
310 310 CTTCCAGGAGGTAC L P G G T	370 AATGGTTCTTCTCA N G S S Q	430 AATAACAAAGCCTA N N K A Y	490 ATCTTTGGCTCAAA I F G S K	550 GGTGACAGATTTGG G D R F G
310 GAG(	370 rctt(	430 AAGC	490 GCT(	550 GAT7
	S	CA A	TGG G	CA CA
310 310 CTTCCAGGAG L P G G	TGG G	TAA	CTT F	TGA D
רניים	AA N	N A.	AT( I	9

300	280	260	240	220
850 860 870 880 890 900 CAGAAAAAGTGTCATGGACATGGGGTATGTAATAGCAATAAGAATTGTCACTGTGAAAAT Q K K C H G H G V C N S N K N C H C E N	AAGATCTGTAGAAACTTCCAGTGTGGTGCTTCTGTTCTG	730 740 750 760 770 780 CTAGGATCAGATCCTGGGATGGTTAACGAAGGCACAAATGTGGTGCTGGA L G S D V P D P G M V N E G T K C G A G	670 680 690 700 710 720 GTGCCTGCTATTATTCAAACGCCTAGTCGAGGCACCCAAATGTTGGGGTGTGGATTTCCAG V P A I I Q T P S R G T K C W G V D F Q	610 620 630 640 650 660 AATGCTTTGTGGGAAGCTTCAGTGTGGAATGTACAGAGATACCTGTATTTGGAATT N A L C G K L Q C E N V Q E I P V F G I
FGA E	Q.	IGC. A	FTT	1992 1992
ري د د	C C	ည်	iGA7 D	TT F
890 TCA(	830 TGAC D	770 ATG1 C	680 690 700 710 72 TCAAACGCCTAGTCGAGGCACCCAAATGTTGGGGTGTGGATTTCCA	650 TGTA V
TTG	ITA	AA;A K	ව්වුව	ACC.
GAA'	SAA'.	CAC	rtg(	GATA I
880 ATAA(	820 TTCT	760 3AAGG	700 AATG	640 AAGA(   E
CAA.	V V	7. CGA. E	CAA.	ACAL 6
TAG S	ITC S	raa N	CAC	rgt. V
LAA.	r PGC	GGT.		SAA?
870 ATGT/ C 1	810 VGAT	750 3ATG	690 rcga R	630 GAG/ E l
3GT/ V	rg T/V	) 999	ragi S	FTG C
) ၁	STG C	ည	ည်င္သ	ر م
860 ACAT H	800 CCAC	740 'AGA1 D	680 AAC( T	620 GCTJ L
TGG	CTT	7CC		AAA K
ICA H	AAA N	rgT' V	ľAT' I	
850 AGTG	790 GTAG	730 CAGA'	670 CTAT	610 TGTG
K AA 8	77 C. C.	7: ATC. S	67 76 7	6] ITT L
3AAJ K	JAT(	£ 1997	GTGCCTGCTATTA:	rgc: A
CA(	AA(	CT	· GT( V	AA7 N

320

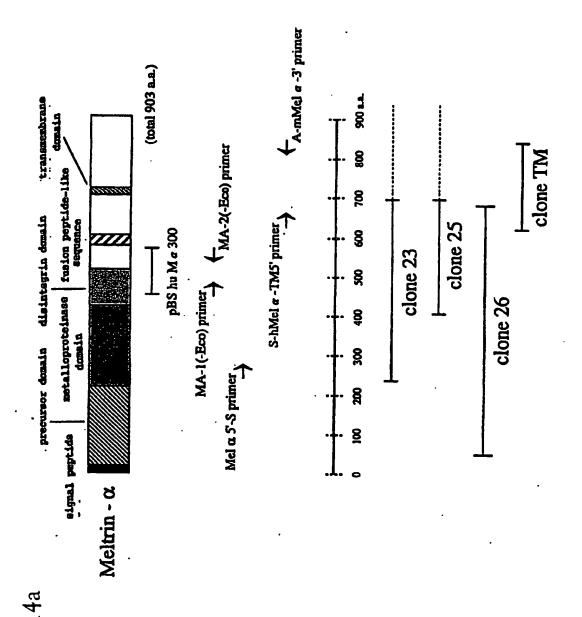
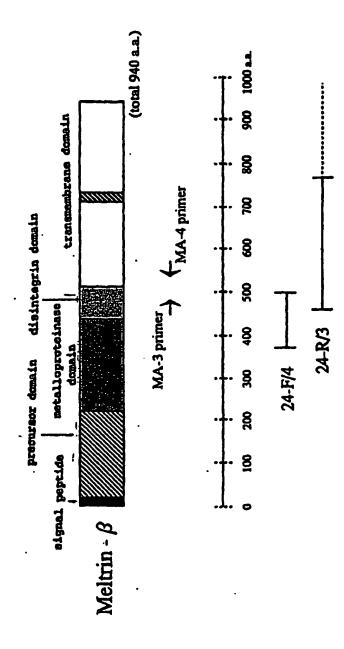


FIG. 14b



# FIG. 15a

09	20	120	40	180	09	240	80	300	100	360	120	420	140
GGGGACCTCTGGATCCCAGTGAAGAGCTTCGACTCCAAGAATCATCCAGAAGTGCTGAAT	z	ATTCGACTACAACGGAAAGCAAAGTAATCATAAATCTGGAAAGAAA	<b>ـ</b> ـا	ATTGCCAGCAGTTTCACGGAAACCCACTATCTGCAAGACGGTACTGATGTCTCCCTCGCT	¥	CGAAATTACACGGGTCACTGTTACTACCATGGACATGTACGGGGATATTCTGATTGAGCA	₹	CTAT	V S L S T C S G L R G L I G F E N E S Y	ecre	V L B P M K S A T N R Y K L F P A K K L	AAAAGCGTCCGGGGATCATGTGGATCACATCACAACACCCCAAACCTCGCTGCAAAGAAT	z
3CT(	_	AGG1	ى	CCT	_	rtc/	S	A A G	S	GAAC	×	AAA	×
AGT(	>	TGA.	<b>5</b>	CTC	S	TGA.	<b>-</b>	TGA	ഥ	GAA	×	TGC	¥
AGA,	E	AAA'	Z	TGT	<b>A</b>	TTC	တ	AAA	z	AGC	¥	ည်	¥
TCC	۵.	AAG	24	TGA	<b>A</b>	ATA	>	TGA	Œ	ည	٥.	CCT	_
TCA	=	499,	<u>च्</u> य	TAC	⊢	999	ပ	GTT	Œ	CTT	(z.	AAA	z
GAA	Z	TCT	_	993	G	ACG.	~	TGG	G	ACT	_	;ACC	۵.
CAA	<b>\</b>	AAA	Z	IAGA	9	TGT	>	lTAT	-	CAA	₩ .	CAC	$\vdash$
CTC	S	CA1	-	lg C.A	0	ACA	×	;ACJ	1	3AT/	<b>X</b>	VCA/	Z
)CG/	9	IGAT	—	TCJ		TG	ၒ	999	9	CAC	<b>~</b>	\TC/	==
CTJ	لت	. ACJ	_	\CT	<b>&gt;</b> -	/22/	Ħ	ÇY	~	CA/	Z	SAC/	Ħ
IGAG	S	AGA1	ਲ	733	==	\CT/	>-	TC	<b>_</b>	CAA(	<b>—</b>	3AT(	တ
LGA/	<b>\</b>	JCA/	<b>≥</b> 4	AAAC	<b>-</b>	STT/	>	CTG(	Ç	STG	~	3TG	၁
CAG	>	VAA(	ဟ	7990	Œ	<b>ACT</b>	ပ	STT(	S	4AA(	S	CAT(	ပ
22	۵.	7999	m	ICA(	T	3TC	×	CGT(	ပ	ľGA.	×	3AT(	S
3GA7	_	IAC(	~	TT	[ ]	990	S	3CA(	$\vdash$	CAA?	æ	999	9
lCT(	₽	rAC/	<b>O</b>	SCAC	S	ACAC	<b>(</b>	CAC	S	AAC	۵.	ည်	<b>∞</b>
ACCI	-	3AC1	<b>_</b>	CCAC	S	ATT/	<b>X</b>	GTC	<b>-</b>	rag/	Œ	ည်	^
7999	9	TTC	~	TTG(	Y	GAA	Z	TCA(	S	TCT	7	A A A (	K S V R G S C G S H H N T P N L A A K N
Ğ	9	V		×		Ö	24	Ġ	>	ی	>	~	¥

## FIG. 15k

48(	16(	54(	18(	90	20(	999	22(	720	240	780	260	840	280
GTGTTTCCACCACCTCTCAGACATGGGCAAGAAGGCATAAAAGAGAGCCCTCAAGGCA	_	AA	TKYVELVIVADNREFQRQGK	GATCTGGAAAAAGTTAAGCAGCGATTAATAGAGATTGCTAATCACGTTGACAAGTTTTAC		AA		J.	CSVSQDPFTSLHEFLDWRKM	A A		<b>Y</b> S	
<b>V</b>	•	SAA	<b>&gt;</b>	E	<b>&gt;-</b>	Ş	$\simeq$	<b>Y</b> 5	Z	ည	ď	ည	S
CA	×	AG	S	GT	Œ	/95	9	SA/	×	E	( <u> </u>	9	ပ
Ž	_	Ş	0	AA	<b>×</b>	AT	=	¥.	<b>~</b>	TA	>	JC.	S
¥C	<b>-</b>	<b>A</b> 66	œ	3AC	0	3AC		ည	-	ïTT	_	AG	~
A G	•	AG		TT		AT(		Ş	<b>Ja.</b> -	999		ည္ဆ	3
3AG	<del>(C)</del>	Ţ	0	9	<b>&gt;</b>	YS.	Z	25	0	25	S	AG.	
AA(	~	G	لته	ζŢ	=	GTG	2=	ICT	_	SAG	S	ည္တ	V
ΓAΑ	<b>×</b>	1GA	드	[AA	Z	ET.	^	II	Œ	ST ST	>	YC	<b>(</b>
CAJ	E .	3	œ	ີວິ	₩	₹Y5	ம	GAA	പ	CTI	_	5	()
991	~	AC	_	II		J.		AT		YC.		ŢĈ.	_
3A.	<u> </u>	$\mathcal{L}$	2.	3	7	S	>	္ဌ	==	ည	0	رچ	×
A A (	∝	AG/	9	VOV	Œ	AGG	S	SCI	_	25	~	3AG	S
ည်	<b>⋖</b>	ည်	¥	MT	-	.CT	<b>^</b>	,AG	S	AA.	Z.	AT(	æ
.TG	æ	51	<b>^</b>	TT/		TT	_	) V	$\leftarrow$	CAC	_	ATC	_
CA	£	170		ζ¥	24	:TG	_	T	_	AT		Ç	
¥G/		757		ತ್ತ	Œ,	2	-	X	Œ	္တ	==	္တ	Δ.
<b>T</b>	<b>~</b>	ĊĊ,	>	ည	0	GA.		ပ္ပ	۵.	AT(	S	ၓ္ဌ	¥
STC	S	SCT	_	ſAA	<b>×</b>	ည်	<b>~</b>	Ç.	9	AA	*	AT	**
ည	۵.	CA(	æ	GT	>	ATI	_	CAĞ	C)	ဒ္ဌ	بہ	ည	(5
SC	_	TG		AA		AC		GT(		E	_	ည္	_
YC:	۵.	TG	<b>&gt;</b>	AA	24	Ę.	2	AA	S	Ş	4	C.	
ည့ <i>်</i>	ച	GTA	>	V S	त्य	ĄČ	-1	lGT	>	Ş	_	SAC	←
TT	Œ,	AA	×	Ę	_	2	RPLNIRIVLVGVEVWNDMDK	JC	S	CI	K L L P R K S H D N A Q L V S G V Y F Q	ACC	G T T I G M A P I M S M C T A D Q S G G
GTG	>	ACT	<b>-</b>	GAT	_	4GA	<u>ب</u>	ည	()	JAG	_	35	· -
	=	_	•					F-4	_	~	, <del>-</del> E,	ت	ن

## FIG. 15c

900	300	960	320	102	34(	108(	36(	114(	38(	120(	40(	1260	420
ATTGTCATGGACCATTCAGACAATCCCCTTGGTGCAGCCGTGACCCTGGCACATGAGCTG		GGCCACAATTTCGGGATGAATCATGACACTGGACAGGGGCTGTAGCTGTCAAATGGCG	A	GTTGAGAAAGGAGGCTGCATCATGAACGCTTCCACCGGTACCCATTTCCCATGGTGTTC	لعرا	AGCAGTTGCAGCAGGACTTGGAGACCAGCCTGGAGAAAGGAATGGGGGTGTGCCTG	_	TTTAACCTGCCGGAAGTCAGGGAGTCTTTCGGGGGCCAGAAGTGTGGGAACAGATTTGTG	^	GAAGAAGGAGGAGTGTGACTGTGGGGAGCCAGAGGAATGTATGAATCGCTGCTGCAAT	Z	GCCACCACCTGTACCCTGAAGCCGGACGCTGTGCGCCACATGGGCTGTGCTGTGAAGAC	_
rgA(	Œ	AAT(	=	3GT(	>	TGC	ပ	TTI	Cz.	TGC	ပ	CGAA	Œ
ACA.	<b>=</b>	TCA.	ð	CAT(	æ	GGT(	Λ	/9Y3	~	crec	ပ	CTGI	ပ
<b>၁</b> 99	IVED HSDNPLGAAVTLAHEL	CTG	G H N F G M N H D T L D R G C S C Q M A	TCC	VEKGGCIMNASTGYPFPMVF	999	5	GAA	FNLPEVRESFGGQKCGNRF	TCG	~	GTG(	ATTCTLKPDAVCAHGLCCED
CCT	<b>-</b>	TAG	တ	ATT	(z.	AAT	æ	TGG	ၒ	GAA	Z	GCT	٦
CGAC	⊢	CTG	ပ	ည	۵.	YAGG	ပ	CTG.	ပ	TAT	æ	TGG	G
ည်သ	>	9999	S	GGTA	>	AGAA	×	\GAA	×	IATG	ပ	ACA.	==
CAG	¥	ACA	~	9	S	TGG/	त्य	/၁၁	0	4GG/	Œ	55	Y
GTG	×	TGG	9	CCA		ည်	-	555	G	CAG/	Œ	IGT(	ပ
ITG	<b>6</b>	CAC	_	CTT	S	CA	S	9	G	ÿ	<u>م</u>	T.	>
$\frac{1}{2}$		CA	<b>F</b>	357	¥	)GA(	<b>—</b>	TT	( <del>2</del> ,	799	田	3	¥
ITC.	۵.	TG/	9	GA	Z	99	Œ	GTC	S	TG	ۍ	799	9
CA/	Z	TCA	×	ČA1	∑i	CTJ	<u>ں</u>	GGA	Œ	CTG	ပ	ည	<u> </u>
,AG.	9	Ç.	Z	CA1	_	755	<u> </u>	CAG	<b>∞</b> :	TGA	0	GAA	<b>5</b>
TTC	တ	GAT	æ	CTG	ပ	GAA	*	AGŢ	<b>^</b>	GTG	ပ	ည	
ည္	=	990	ය	AGG	ပ	CAG	8	GGA	त्य	GGA	œ	TAC	⊣
<b>4</b> 25	9	TTI	Œ	AGG	ပ	CAG	S	ည္ဟ	۵,	AGA	ഥ	CTG	ပ
CAT 	<b>=</b> 6	CAA	z	GAA	×	TTG	ပ	SCI	_	AGG	ပ	CAC	←
Tet	>	CCA	æ	TGA	( <del>1</del> )	CAG	S	TAA	Z <sub>.</sub>	AGA	Œ	CAC	<b>-</b>
AT	<b>-</b>	99	ပ	GT	<b>&gt;</b>	AG	S	1	Œ	QA.	(z)	ij	¥

1320	1380	1440	1500	1560 520	1620	1680 560
TGCCAGCTGAAGGAACAGCGTGCAGGACTCCAGCAACTCCTGTGACCTCCCA C Q L K P A G T A C R D S S N S C D L P	GAGTTCTGCACAGGGCCAGCCTCACTGCCCAACGTGTACCTGCACGATGGGCAC R F C T G A S P H C. P A N V Y L H D G H	TCATGTCAGGATGTGGCTACTGCTACATGGCATCTGCCAGACTCACGAGCAGCAG	TGTGTCACGCTCTGGGACCAGGTGCTAAACCTGCCCTGGGATCTGCTTTGAGAGGTC C V T L W G P G A K P A P G I C F E R V	AAA (	CA	990
J.	99	<b>V</b>	GA(	22	99	ည္မ
227	TG G	) )	GA R	TG A	ပ္သံ 🕿	AG.
Τζ. D	ල	ර්ථ ක	ည်အ	CTI F	S	99
င်းင	SCA H	TCA	CTT	rrc s	76C.	CA/
CTC S	CCT	GAC	CTG C	GAG. S	<b>\</b> 663	)CA(
CAA	GTA Y	Q CCA	GA T	GAA( K	466. 6	SCT L
CAG S	CGT V	CTC C	TGG G	S S	ľCA.	) ) 
CTC S	CAA	CAT	) 200	AGT(	3 <b>T</b> G	CAT(
0	AGC A	TGG G	TGC A	CAA,	OCA(	AAA( N
CAG	ည္သ	CAA	ACC P	766 6	AAT(	AAC, T
ST3 C	CT C	CTA Y	TAA K	CTG	AAA. K	AGA. E
A G C	TCA H	CTG C	TGC A	CAA	766 G	CAT
AAC	ည္သ	CTA Y	AGG G	TGG	ATG C	S
AGG G	CAG S	9 99	ACC P	TTA Y	TAA. K	CGT'
TGC	099 V	GGA D	999	ر ۲	TGC.	rgc(
်ည် ႕	AGG G	TGT V	CTG ■	TGA D	AGA'	CAA'
, K	CAC	GGA D	GCT L	AGG	GAG	ľAC T
125	CTC C	TCA Q	CAC	TGC A	GAT	ည်ပ
TGCCAGCTGAAGGAACAGCGTGCAGGACTCCAGCAACTCCTGTGACCTCCC	GAGTTCTGCACAGGGCCAGCCTCACTGCCCAGCCAACGTGTACCTGCACGATGGGC B    F    C    T    G    A    S    P    H    C. P    A    N    V    Y    L    H    D    G    H	ATG C	TGT V	AATTCTGCAGGTGATCCTTATGGCAACTGTGGCAAAGTCTCGAAGAGTTCCTTTGCCAAA N S A G D P Y G N C G K V S K S S F A K	TGCGAGATGAGATGCTAAATGTGGAAAAATCCAGTGTCAAGGAGGTGCCAGCCGGCCA C E M R D A K C G K 1 Q C Q G G A S R P	GTCATTGGTACCAATGCCGTTTCCATAGAAACAACATCCCCCTGCAGCAAGGAGGCCGG V I G T N A V S I E T N I P L Q Q G G R
16 ဂ	S B	JC. S	7G'	AA. N	7G, C	GT( V

## FIG. 15e

1740	280	1800	1860	1920 640	1980	2040	2100
<b>ATTCTGTGCCGGGGACCCACGTGTACTTGGGCGATGACATGCCGGACCCAGGGCTTGTG</b>	>	CTTGCAGGCACAAAGTGTGCAGAAAATCTGCCTGAATCGTCAATGTCAAAATATT L A G T K C A D G K I C L N R Q C Q N I	AGTGTCTTTGGGGTTCACGAGTGCAATGCAGTGCCACGGCAGAGGGGTGTGCAACAAC S V F G V H E C A M Q C H G R G V C N N	AGGAAGAACTGCCACTGGGCACCTCCCTTCTGTGACAGTTTGGCTFT R K N C H C · E A H W A P P F C D K F G P	GGAGGAAGCACAGCGCCCCATCCGGCAAGCAAGCAAGCAGGAAGCTGCAGAG G G S T D S G P I R Q A E A R Q E A A E	TCCAACAGGGAGCGCCAGGAGCCCGTGGGATCGCAGGAGCATGCGTCTACT S N R B R G Q G Q B P V G S Q B H A S T	GCCTCACTGACACTCATCTGAGCCCTCCCATGACATGGAGACCGTGACCAGTGCTGCTGC A S L T L I *
CII		AAT N	AAC N	999	V QC/	TCT S	าัธ
995	( )	CAA	ည္သ	rtt F	GCT A	SCG A	J
Š	_	E)	;TG.	.yv.	, AA	ATC	;AG;
ACC	<u>a.</u>	AAT	099	ACA	AGG P	AGC F	¥CC
99	0	)]C	3AG G	orc O	) (0	757 B	3TG
ည္ဟ	<b>a</b>	77. R	CA(	ည်	AA( R	) ()	ည
CAT	<b>=</b>	GAA N	999	CTI F	A G C	ATC S	AGA
GA	9	SCT L	ξĊ. ₩	ეე _	AGA E	99 <u>9</u>	76G
,GA	9	ည်း	ည်း	20 4	V V	Y V	CA.
99	Ç	ATC 1	OV O	SC.A	¥ o	) )	TGA
TTG	_3	A A A K	A T G	76G	200 8	GAG	CCA
'AC		, GA,	SC Z	YYC	) []	AG.	). 
<u> </u>	•	AT6	6T6	33	22 _	25	ည္တ
ACG	<b>&gt;</b>	CAG	AGT	AGG A	ည္	AGG G	GAG
23	=	TG A	9) (B)	ည်း	၁	ည်င	ĬĊŢ*
SA(	<b>-</b>	SI C	TC/ m	ည်သ	CAC	ည္တိဗ	CA7
999	ၒ	AAA K	GGT V	SC ₩	AGA D	ეე ™	ACT L
Š	<b>~</b>	JAC	ည် ဝ	ည်	JAC	SGA B	3AC T
736	ပ	999	F	N N	S S	.AG(	CT
CTO		SC.A	GTC V	A A G K	<b>7</b> 99	N N	TC/ S
ATT	I L C R G T H V · Y L G D D M P D P G L V	CTTGCAGGCACAAAGTGTGCAGAAAATCTGCCTGAATGTCAATGTCAAAAT L A G T K C A D G K I C L N R Q C Q N	AGTGTCTTTGGGGTTCACGAGTGTGCAGTGCCACGGCAGAGGGGTGTGCAAC S V P G V H E C A M Q C H G R G V C N	AGGAAGAACTGCCACTGCGCCACCTCCCTTCTGTGACAAGTTTGGCT R K N C H C E A H W A P P F C D K F G F	GGAGGAAGCACAGCGCCCCATCCGGCAAGCAAGCAAGGCAGGAAGCTGCAGA G G S T D S G P I R Q A E A R Q E A A E	TCCAACAGGGAGCGCCAGGAGCCCGTGGGATCGCAGGAGCATGCGTCTA( S n r e r g q g q e p v g s q e h a s t	GCCTCACTGACACTCATCT A S L T L I *
-							

	ACTGCATAAAATAGAGTGCATCCCGCCC 2848
2820	AGTAAGAATGTTAAAAAGTGAAAACAATGTAAGAGCCTAACTCCATCCCGTGGCCATT
2760	AGGTTTGAGGGTTTGCAGAAAGCCAGGGAACCCACAGAGTCACCAACCCTTCATTTAACA
2700	TAGGAGAAAGGGCGGTGAACTCTGGCTCTTTGCTGTGGACATGCGTGACCAGCAGTACTC
2640	CACCCATTCCATCTCCATCCAAGCAAACTGAATGGCATTGAAACAAAC
2580	TTTAGCATTTATTATATGAAAATAGCAGGGTTTTAGTTTTTAATTTATCAGAGACCCTGC
2520	CTCAGCCTTGGCAGCCCTGATGACTGGTCTCTGGCTGCAACTTAATGCTCTGATATGGCT
2460	CCCCACAGCAGTGGGGGAGAAGCAAGGGTTGGGCCCAGTGTCCCCTTTCCCCAGTGACAC
2400	GCAGGCCCCAGCCTGCAGCAAGGAGGAAGACTCAAAAGTCTGGCCTTTCACTGAGC
2340	CTGCTAAAACATGGACATGCTTCAGTGCTGCTCCTGAGAGAGTAGCAGGTTACCACTCTG
2280	CCTACCAGGCACGTCTGCAGAACAGTGCAAGGAAGGGCAGCGACTTCCTGGTTGAGCTT
2220	CCATCGTTTCCATGACAACAGACACACAGTTCTCGGGGCTCAGGAGGGGAAGTCCAG
2160	AGAGGAGGTCACGCGTCCCCAAGGCCTCCTGTGACTGGCAGCATTGACTCTGTGGCTTTG

 3

 3

 4

 3

 3

 4

 4

 5

 6

 6

 6

 7

 8

 8

 9

 9

 10

 10

 10

 10

 10

 10

 10

 10

 10

 10

 10

 10

 10

 10

 10

 10

 10

 10

 10

 10

 10

 10

 10

 10

 10

 10

 10

 10

 10

 10

 10

 10

 10

 10

 10

 10

 10

 10

 10

 10

 10

 10

 10

 10

 10

 10

 10

 10

 10

 10

 <td

					•	
Þ .,	AAG K	GAG TTC B F	년	9 0	£ _	
C 1	H 34	<b>6</b> 1	C 2	ည	F)	
8 4	ဠပ	95 195	8 -	ဋ္ဌ	₹ z	•
AAT N	ACC CTG AGG CCG CGG TGT CCT CAC CCC TCC TCC TCC CAC TGT T L R P G A E C A H G S C C H Q C	CAG COCC AGG CAG TGT GAC CTC COCC	1 ABC	200	) 	
ဥ္သ	S <b>≖</b>	ر ا	5	<b>2</b> 80	A&g ™	
ည္သ	ည္သ	) 28 2	GAT	ATG TGC CTC ACC TAC CAG	S S S	
8 2	ည္ဆင့္သ	C E	ATC	TAC	E .	
N AC	B <sup>'</sup> 8	၌ တ	ည် ဇ	<b>₩</b>	ည္ဆ	<b>%</b> &
AAC N	ည်	A 36	Y Y	CC CC	ည ၂	394 131
C []	E CC	<b>2</b> 0 <b>√</b> .	77C F	ည္ညည	D G	
<b>₩</b> ₩	GCT A	95 <i>0</i>	AAC	ATC	الم الم	CAC D
<b>9</b> 8	16T C	S E	ACC T	<b>3</b> 6	<b>3</b> €	AAG K
AS ES	EAG E	OCT GOG ACC CTG TGC CGC CAG P G T L C R B	<b>P</b> . CC	CAG GCC TAC TGC TAC AAC GGC Q A Y C Y N G	الم الم	<b>8</b> 0
AR E	900 V	ည္ဆ	ည္ဆ	TAC Y	ই ≃	15T C
<b>E</b> 5	<b>35</b>	<u>5</u>	S E	ည္ည	გ •	AAC
<b>国</b> 。	<b>8</b>	ABC T	8 4	7 Y	<b>8</b> 50	<del>§</del> 5
E EXT	AGG R	9	S	<b>3</b> 5 ▲	8 ~	E 4
C E	را د	الم الم	AAG K	900	. 88 2	7 7
95 B	ACC T	A A	ည္တ	ည္တ	70G ₩	D CEC
GGG GAA GAG TGT GAT TGT GGA GAA GAG GAA TGT AAC AAC CCC TGC TGC AAT GCC TCT G E E C D C G E E E E C N N P C C N A S	151 C	CTG TTG CCT CCTC CCTG TCC CCC CAG CCC AGG CAG TCT CAC CTC CCCC	TOT ACTS ONCE AND TOTO COCK ONCE AND TOTAL CHICAGO AND GOAT ONCE TOTAL ON TOTAL ON THE ONE ONE ONE ON THE ONE ON THE ONE ON THE ONE ONE ONE ONE ONE ONE ONE ONE ONE ON	GAG GOC CAG GCC TAC TGC TAC GCC ATG TGC CTC ACC TAC CAG GAG CAG TGC CAG E G G Q A Y C Y N G M C L T Y Q E Q C Q	CAG CTG TOG GGA CCC CCC CCC CCC CCC TCC CTC CAG AAG GTG AAT GTG Q L W G P G A R P A P D L C F E K V N V	GCA GGA GAC ACC TITT GGA AAC TGT GGA AAG GAC A A G D T F G N C G K D
<b>38</b> 5	AAT TGT A	) 13	151 C	SAG E	9 0	<b>8</b> 4

# FIG. 17a

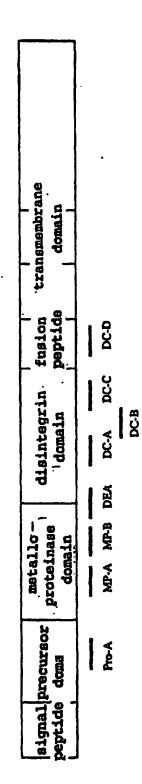
60 20	120	180	240	300	360 120	420
GCT L	GAT	orc C	TGGAGAAGAAGGAATGTAACCAACCCCTGCTGCATGCCTCAATTGTACCCTGAGGCC	3AC T	ر ا ا	STA Y
GGA( E	CAG R	rga( D	3AG R	, 9	STC) S	7) V
SAG R	CAC	3TG C	) 1	7 P	SAA( K	CAC
CAG(	AGA( D	AGA( E	racc T	36CJ A	9	) ) )
CAAC	30C/ P	B. B.	rtg: c	TT( L	racc T	999
ATG( C	CAT(		LAA'	J L	).TG] C	rca( B
rgg, G	CAAC	AGA.	S	r K	STT( F	21G7 C
CAA'	CTC S	GGA. B	rgc(	3 <b>T</b> G)	SGA(	)) 
GTT F	ľČŢ L	lCT L	CAA.	y လ	) 	rac(
AGT(	GTG	STA'. Y	CTC C	CCA(	J T	[66] G
CAA. K	AAT(		်ာင်	င်းရ င	G D	GA1
TCC P	766, 6	GAA( N	) )	CTG(	31G C	SAT(
CTT F	TGG	766 6	CAA	CTC S	SCA(	OCA(
CCC P	AGG	GTG	TAA	999	CAG R	CTA( Y
GCA H	GTC S	GAG R	ATG C	TCA H	SGC.	CTT F
TGG G	OCA O	S R	GGA E	TGC A	GCA Q	CAA
CAC	TCT	AGG G	AGA E	GTG C	CGA( E	TAC
TGC A	GTA.	766 G	AGA. E	GGA(	) 2002 8	ည်႕
AGC · A	CAG R	GTA. Y	AGA E	<b>∨</b>	ST S	CTG
CGGAGCTGCCACTGGCACCCTTTCCCAAAGTGTTCAATGGATGCAACAGGAGGGAG	GGACAGGTATCTGCAGTCGTGGAATGTGTCTCTCCAACATGCCAGACACCAGGAT D R Y L Q S G G G M C L S N M P D T R M	GTTGTATGGAGGCCGGAGGTGTGGGGAAGATGGGGAAGAGTGTGACTG L Y G G R R C G N G Y L B D G E E C D C	TGG. G	GGGGGGGGAGTGTGCTCCTGCTGCCACCAGTGTAAGCTGTTGGCTCCTGGGAC	CCTGTGCCGCGGCCAGGCAGTGTGACCTCCCGGAGTTCTGTACGGGCAAGTCTCC L C R E Q A R Q C D L P E F C T G K S P	CCACTGCCCTACCAACTTCTACCAGATGGATGGTACCCCCTGTGAGGGCGGCCAGGCCTA H C P T N F Y Q M D G T P C E G G Q A Y

480	540 180	600 200	. 660	720	780	840 280
CTGCTACAACGCCATGTGCCTCACCAGGAGCAGTGCCAGCAGCTGTGGGGACCCGG	AGCCCGACCTGCCTCTGCTTCGAGAAGGTGAATGTGGCAGGAGACACCTTTGG A R P A P D L C F E K V N V A G D T F G	STG C	I	2 2 2 3	T. C	S B
22 4	TTI	AAG K		TAC	A A G K	CT
ge P	ည္	926	TG.	TC	22.	AA,
9	AC.	ATC	990	ACG V	SAA	rtg
rg.	JAG D	AG D	V V	ဘွဲ့ 🗷	ີ່ ໄດ້ ດ	CI F
<u>පු</u>	Y C	GA(	CAA	CAC T	SAC	TT.
ည် လ	) A	ZAT ⊯	S.	ပ္ပံ ပ	A TC	S S
ပ္လ တ	`````` V	N N	GA(	) ) ) (	GTG V	ACC T
). C	AA7 N	) ) )	CTG L	76C C	CTG L	AAC
CA(C	GTG V	A A G	22	O O	ე ე	466.
GAG	AAG K	466 ?	990~	TC	νς. CV	25.
AG.	AG.	AC.	22	AG.	ACC P	AGI
22	ည် အ	IAC H	991 V	ည္တဲ့တ		ည္
CT.	CIT F	TG/ B	16.A E	GAG R	SCT	99 <u>5</u>
CAC	STS C	ე ე	S	ပ္သိ ပ	A T (	GA(
1 1	T)	SAA?	s S	N N	Q D	CTI
STG C	<u>[G</u> A(	ATC H	CAC O	ATC #	5	76C C
TGCTACAACGCCATGTGCCTCACCTACCAGGAGCAGTGCCAGCTGTGGGGACCCG(	)   	AAACTGTGGAAAGGACATGGTGAACACAGGAAGTGCAACATGAGATGCGAAGTG N C G K D M N G E H R K C N M R D A K C	TGGGAAGATCCAGTGTCAGAGCTCTGAGGCCCTGGAGTCCAACGCGGTGCCCAT G K I Q C Q S S E A R P L E S N A V P I	TGACACCACTATCATCATGGGAGGCAGTCCAGTGCCGGGGCACCCACGTCTACCG  D T T I M N G R Q I Q C R G T H V Y R	AGGTCCTGAGGAGGGTGACATGCTGGACCCAGGGCTGGTGATGACTGGAACCAAGTG G P E E E G D M L D P G L V M T G T K C	TGGCTACAACCATATTTGCCTTGAGGGCAGTGCAGGAACACCTCCTTCTTTGAAACTGA G Y N H I C L E G Q C R N T S F F E T E
ည္တ	) V	IAA(	CAC Q	ATC	GAG E	CAT
N N	ار 1	GGA G	ATC I	ACT T	GAG E	AAC
TAC Y	CGA R	<b>TGT</b> C	AAG	200	CT	rac.
] ]	225	NAC.	99	AC	FE	[] []
ິວ	) Y	A &	76 0	75 U	AG G	រិក ភ

# FIG. 17c

300	960 320	1020 340	.1080	1140 380	1183 394
AGGCTGTGGGAAGAGTGCAATGGCCATGGGGTCTGTAACAACAACAGAACTGCCACTG	CCTGCCGGGCTGGGCCCCCTTCTGCAACACACGGGGCCACGGGGGCAGTATCGACAG	TGGGCCTATGCCCCCTGAGAGTGTGGTCCTGTGGTAGCTGGTGGTGGTGGCCATCTT	CTAGGCCA	TTCAGGGT	
G C G K K C N G H G V C N N N Q N C H C	L P G W A P P F C N T P G H G G S I D S	G P M P P E S V G P V V A G V L V A I L	L G Q	F R V	
AACCAGAAC	GGGGGCAG1	GTGTTGGT@	AACAACAAA	AGTTGTCCC	• .
N Q N	G G S	V L V	N N K	S C P	
TAACAAC	CGGCCAC	AGCTGGA	CAGACAGA	ACAGTTC/ Q F S	TTTCAAG F K
GGGGTCTG	AACACACC	CCTGTGGT	TACTGCTG	CTGAGGCA	AACCCAAC
G V C	N T P	P V V	Y C C	L R Q	N P T
ATGGCCAT	ccticico	GTGTGGGT	IGATGTAC	S K	STCATGCC
G H	F c	V G	M Y		H A
AAGTGCA.	GCCCGC	CCTGAGA(	CTCATGC;	GCTCTCC(	GGGACTG(
K C N		P E S	L M L	A L P	G T G
IGGCTGTGGGAAGAGTGCAATGGCCATGGGGTCTGTAACAACAACAGAACTGCCACT( G C G K K C N G H G V C N N N Q N C H C	CTGCCGGGCTGGGCCCCCTTCTGCAACACACGGGGCCACGGGGGGAGTATCGACA	GGGCCTATGCCCCTGAGAGTGTGGTCCTGTGGTAGCTGGTGGTGGTGGCCATCT	GGTGCTGGCGGTCCTCATGTACTACTGCTGCAGACAGAACAACTAGGCCA	ACTCAAGCCCTCACCTTCCAAGCTGAGGCAACAGTTCAGTTGTCCCTTCAGGGT L K P S A L P S K L R Q Q F S C P F R V	TTCTCAGAACAGGGACTGGTCATGCCAACCCAACTTTCAAG S Q N S G T G H A N P T F K
AGGC.	CCTG	T6660	66TG( V	ACTCA L K	TTCTC S Q

 ${\rm FIG.}\ 18a$  Peptides used for the preparation of monoclonal antibody



 ${
m FIG.}\ 18{
m b}$  Peptide sequences used for the preparation of monoclonal antibody

No.	Name .	Sequence (N-terminal, C-terminal)
1	Pro-A	T T D S Y K L V P A E S M T N I C
7	MP-A	ADNREFQRQGKDLEKVKC
3	£P-8	FTRLHEFLDWRKIKC
4	DC-A	QLKPPGTACRGS SNSC
5	BC-B	GTACRGSSNSCDLPEFC
9	)-JQ	GKDSKSAFAKCELRDAKC
7	ด-วด	QGGASRPV I GTNAVS I ETN I C
8	DE-A	LFNLPEVKQAFGGRKC

FIG. 19 Western blotting with anti-Meltrin monoclonal antibodies

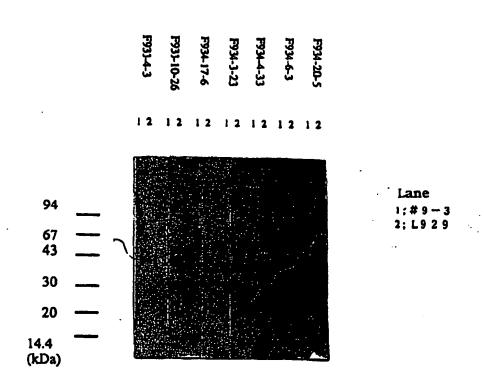
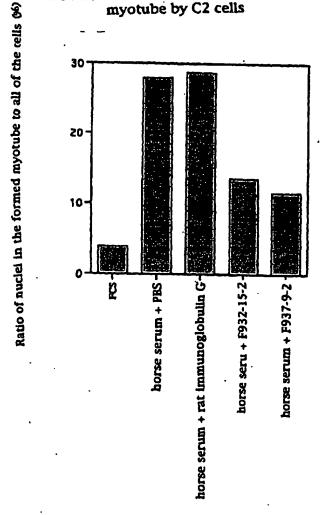


FIG.~20 Effects of anti-mouse Meltrin antibodies on the formation of myotube by C2 cells



Ratio of nuclei in the formed myotube to all of the cells (%)

FIG. 21 Effects of anti-mouse Meltrin antibodies on the formation of pit (bone-resorption area) in mouse unfractionated bone cells

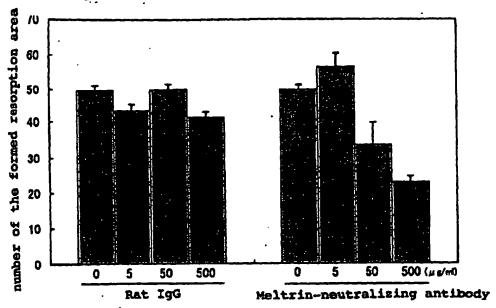
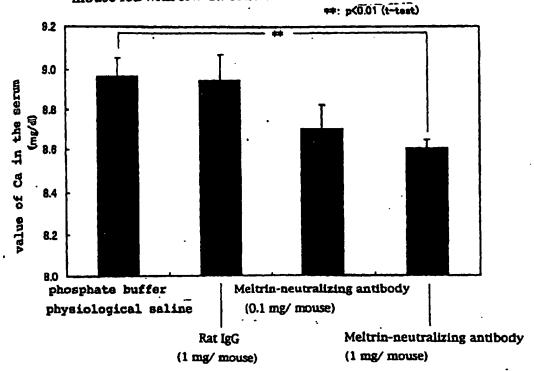


FIG. 22 Effects of anti-mouse Meltrin antibodies on the serum Ca values of the mouse fed with low Ca-content feed



### FIG. 23*s*

9 8 **GCACAAAGTGTGCAGATGGAAAATCTGCCTGAATCGTCAATGTCAAAATATTAGTGTCT** ဟ z 0 ပ 0 2 z \_ ပ -**~** G 9 ¥ ပ **노** 

120 40 TTGGGGTTCACGAGTGTGCAATGCAGTGCCACGGCGGGGGGTGTGCAACAACAAGA 4 2 Z Z ပ > S 24 Ç H O ð × 4 ပ 四 Ħ ය

8 180 ACTGCCACTGCGAGGCCCACTGGCCACCTCCCTTCTGTGACAAGTTTGGCTTTGGAGGAA ى G Ľ G بح) **M** \_ ပ ÇŦ. . Д ۵. ~ **}=** H ¥ Œ ပ ပ

240 8 GCACAGACAGCGCCCATCCGGCAAGCAGATAACCAAGGTTTAACCATAGGAATTCTGG ပ **(** \_ G ð Z A D 0 2 م ပ S

300 100 IGACCATCCTGTCTTCTTGCTCCCCCATTTGTCGTTTATCTCCAAAAGGAAGACCTTGA \_ ⊢ ¥ 24 × \_ > **>** > بع) ය ¥ ~ **ــ** \_ ပ \_

360 120 S TACGACTGCTGTTTACAAATAAGAAGACCACCATTGAAAAACTAAGGTGTGTGCGCCCTT \_ 2 > ပ 2 \_ M Œ ⊣ 7 **×** Z **(** بعا ٦.

140 COCGECCACCCATGCTTCCAACCTGTCAGGCTCACCTCGGCCACCTTGGAAAAGGCC ပ 7 S \_ ပ \_\_ = 4 0 ပ ۵. Ø لعا ပ 2 ۵. ۵. 2

### FIG. 23b

160 TGATGAGGAAGCCCCAGATTCCTACCCACACCAACCAATCCCAGGAGATTGCTCCAGT 24 24 ۵. Z 9 **×** ۵. ۵. **>** ဟ **A** ۵. ۵. **×** 24

180 540 GTCAGAATGTTGACATCAGCAGCCCTCAACGCCCTGAATGTCCCTCAGCCCCAGTCAA S 0 م 0 ۵. > Z د ය z u ۵. 24 S 0 Z

CTCAGCGAGTGCTTCCTCCCACCGGGCTCCCAGGCTCCCTGCCAGAC

8 8 8

> CCCTGCCAAGCCTGCACTTA 624 L P A K P A L 207

## FIG. 24a

60	120	180	240	300	360 120	420
SCT L	SAT	CTC C	)  -  -	AC T	္သ သ ရ.	TA Y
E G	AG.	GAC D	AGC R	999	rc1 S	22.
99 ~	22	îgr )	TC	ຸນ	AG	V G V
799	AC.	AG1	22	CTC P	GCA K	ည္ဌ
ACA R	OYC	IAG E	T	ŞŞ.	ပ္တဲ့ ပ	క్ర ఆ
Z Z	ర్ల	β	TT C	GTT L	TAC T	995
ATG	CAT	TGG G	TAA N	1 1	CTG	TGA( E
7760 6	CAA N	AGA D	CTC S	TAA K	GTT	CTG
CAA	CTC S	A E	TGC A	GTG C	SGA B	ည်႕
GT1 F	17. 1	TCT L	CAA	ည် တ	ပ္သိ	rac T
AGT	6T6	GTA Y	CTG C	CCA(	CCT L	ည်
CAA	AAT	<sub>ອ</sub>	CTG	CTG	IGA( D	GA1
TCC	TGG G	GAA	ეე -	င်းငွဲ	316. C	3AT(
CTT F	TGG G	າດ ດ	CAA	S S	O O	o O
ည	AGG G	ore c	IAA. N	်ပ္သိပ္	CAG( R	TA( Y
GCA H	GTC S	GAG R	ATG C	CA(	) V	TT( F
TGG G	QCA O	<u>ين</u> 25. هـ	GGA.	] ₽	) O	N N
CAC	TCT L	A G.G.	AGA( E	31G C	B G	T
TGC	GTA Y	TGG	AGA E	GGA( E	<u>کي</u> ۳	)    -
AGC A	CAG R	GTA Y	AGA. E	S A	STG C	ည်
CGGAGCTGCCACTGCCATTCCCAAAGTGTTCAATGGATGCAACAGGGAGGCAGCT G A A T G H P F P K V F N G C N R R E L	GGACAGGTATCTGCAGTCAGGTGGAATGTGTCTCTCCAACATGCCAGACACCAGGAT D R Y L Q S G G G M C L S N M P D T R M	GTTGTATGGAGGCCGGAGGTGTGGGGAAGATGGGGAAGAGTGTGACTG L Y G G R R C G N G Y L E D G E E C D C	TGGAGAAGAAGAATGTAACAACCCCTGCTGCAATGCCTCTAATTGTACCCTGAGGCC	GGGGGCGGAGTGTCACGGCTCCTGCTGCACCAGTGTAAGCTGTTGGCTCCTGGGAC	CCTGTGCCGCGGCCAGGCAGTGTGACCTCCCGGAGTTCTGTACGGGCAAGTCTCC L C R E Q A R Q C D L P E F C T G K S P	CCACTGCCCTACCAACTTCTACCAGATGGTACCCCCTGTGAGGGGGGGCGAGGCCTA

480	160	540	180	600	200	099	220	720	240	780	260	840	
CTGCTACAACGGCATGTGCCTCACCTACCAGGAGCAGTGCCAGCAGCTGTGGGGGACCCGG		557	ARPAPDLCFEKVNVAGDTFG	TG	၁	AT	_	9	~	TG	ပ	V.	: : :
ACC	۵,	CIT	נבי	3AA(	*	ပ္ပ	۵.	TA(	<b>&gt;</b> -	:AAG	<b>×</b>	ACT	
) 995	ပ	CACC	<b>—</b>	ည	¥	SCTC.	>	) CTO	>	ACC	⊣	GAA	Cz.
GTG	<b>}=</b>	AGA(	0	VCA7	<b>A</b>	25	V	CAC	×	CGA	G	TTT	
GCT	CYNGMCLTYQEQCQQLTFGPG	AGG,	S	AAACTGTGGAAAGGACATGATGGTGAACACAGGAAGTGCAACATGAGAGATGCGAAGTG	<b>∝</b>	TGGGAAGATCCAGTGTCAGAGCTCTGAGGCCCGGCCCCTGGAGTCCAACGCGGTGCCCAT	z	TGACACCACTATCATGAATGGGAGGCAGATCCAGTGCCGGGGCACCCACGTCTACCG	DTTIIMGRQIQCRGTHVY	AGGTCCTGAGGAGGAGGTGACATGCTGGACCCAGGGCTGGTGATGACTGGAACCAAGTG	G P E E E G D M L D P G L V M T G T K C	TGGCTACAACCATATTTGCCTTGAGGGGCAGTGCAGGAACACCTCCTTCTTTGAAACTGA	<u></u>
GCA	ð	299	¥	CAT(	Ħ	3TC	S	3660	G	ATC	7	TCC	U
CCA	ď	TGT	<b>&gt;</b>	CAA(	Z	3GA(	ല	95	~	GTG	<b>\</b>	YCC	€
GTG	ပ	GAA	Z	GTG(	ပ	SCT	ے	)TG	ပ	SCTO	-1	AAC	Z
SC.A	ø	GGT	<b>^</b>	GAA	<b>×</b>	ည္ပ	هـ	Ç	ď	)55	5	;AGC	. 🕰
GGA	Œ	GAA	<b>×</b> ,	CAG	~	900	~	SAT(	-	$\tilde{\mathcal{C}}$	۵,	)TGC	ن
CCA	ď	CGA	ट्य	ACA	Ħ	၁၅၅	Ä	SCA	œ	3GA(	9	)CA(	0
CTA	<b>&gt;</b>	CTT	لت	TGA	ഥ	TGA	ద	SAG(	<b>~</b>	SCT	<b>_</b>	99	G
CAC	T	CTG	ပ	TGG	ဗ	CTC	S	1991	5	CAT(	×	JGA(	Œ
CCT	7	CCI		GAA	Z	GAG	S	GAA.	z	rgA(	Ω	ΣΞ	۔۔
GTG	ပ	TGA	9	CAT	=	TCA	0	CAT(	<b>=</b>	3663	ပ	776	ပ
CAT	æ	ည	<b>a</b>	QCA.	Q	GTG	ပ	CAT	<b>—</b>	3GA(	œ	[AT]	
993	ၒ	TGC	¥	AAA	×	CCA	ď	IAT(	-	3GA(	田	ZY.	H
CAA	2	ACC	م	TGG	ပ	GAT	-	CAC	<b>—</b>	lGA(	ഥ	:AAC	z
CTA	>	ည	<b>~</b>	CTG	ပ	GAA	×	CAC	<b>(</b>	CC	۵.	TA(	~
CTG	ပ	AGC	¥	AAA	Z	TGĞ	<b>ڻ</b> .	<b>TGA</b>	9	1994	G	)     	ၒ

### FIG. 24c

AGGCTGTGGGAAGAGTGCAATGGCCATGGGGTCTGTAACAACCAGAACTGCCACTG 900 G C G K K C N G H G V C N N N Q N C H C 300	G	0.25 2 0 1 2 2.0	<b>f</b> ⊷			
TAACAACAACO N N N (	) 9 H 9	AGCTGGAGTG1 A G V I	CAGACAGAACI R Q N N	ACAGTTCAGT1 Q F S C	TTTCAAGCCGG F K P E	CCACGGCCACA
CATGGGGTCTG H G V C	TGCAACACACC	sercereres F V V	TACTACTGCTG	AGCTGAGGCA [ L R Q	CCAACCCAAC	GCCACCAATT:
NGTGCAATGGC	CCCCCCTTC	TGAGAGTGTG( E S V (	CATGCTGATG1 M L M 1	TCTCCCTTCCA L P S R	GACTGGTCATG T G H A	CCATGACAAGG
GCTGTGGGAAGAA C G K K	CTGCCGGGCTGGGCCCCCTTCTGCAACACACGGGGCCACGGGGCAGTATCGACA	SECCTATECCCCC	recteccestect L A V L	CCAAGCCCTCAGC K P S A	TCAGAACAGCGG Q N S G	CCACAGCCCACCACCATGACAAGGGCCACCAATTCCACGGCCACACCCTCCTCCACTC

TGGGGACGACCCGGATCCTCACTGAGCTGACCACACAGCCACTACAACTGCAGCCACTG	132
6 D D P D P H *	42,
GAICCACGCCACCTGICCICCACCCCAGGGACCACCIGGAICCICACAGAGCCGAGCA	138
CTATAGCCACCGTGATGGTGCCCACCGGTTCCACGGCCACCGCCTCCTCCACTCTGGGAA	144
CAGCTCACACCCCCAAAGTGGTGACCACCATGGCCACTATGCCCACAGCCACTGCCTCCA	150
CGGTTCCCAGCTCCACCGTGGGGACCACCCGCACCCTGCAGTGCTCCCCAGCAGCC	156
TGCCAACCTTCAGCGTGTCCACTGTCCTCCTCAGTCCTCACCACCCTGAGACCCACTG	162
GCTTCCCCAGCTCCCACTTCTCTACTCCCTGCTTCTGCAGGCCATTTGGACAGTTTTTCT	168
CGCCCGGGGAAGTCATCTACAATAAGACCGACCGAGCCGGCTGCCATTTCTACGCAGTGT	174
GCAATCAGCACTGTGACATTGACCGCTTCCAGGGCGCCTGTCCCACCTCCCCACCGCCAG	180
TGTCCTCCGCCCCCTGTCCTCGCCCTCCCTGGCTGTGACAATGCCATCCCTC	186
TCCGGCAGGTGAATGAGACCTGGACCTGGAGAACTGCACGGTGGCCAGGTGCGTGGGTG	192
ACAACCGTGTCGTCCTGCTGGACCCAAAGCCTGTGGCCAACGTCACCTGCGTGAACAAGC	198
ACCTGCCCATCAAAGTGTCGGACCCGAGCCAGCCTGTGACTTCCACTATGAGTGCGAGT	204
GCATCTGCAGCATGTGGGGGGGGGCTCCCACTATTCCACCTTTGACGGCACCTCTTACACCT	210
TCCGGGGCAACTGCACCTATGTCCTCATGAGAGATCCATGCACGCTTTGGGAATCTCA	216
GCCTCTACCTGGACAACCACTACTGCACGGCCTCTGCCACTGCCGCTGCCGCCGCTGCC	222
CCCGCGCCCTCAGCATCCACTACAAGTCCATGGATATCGTCCTCACTGTCATGGTGC	228
ATGGGAAGGAGGGCCTGATCCTGTTTGACCAAATTCCGGTGAGCAGCGGTTTCAGCA	234

FIG. 24e	
AGAACGCCGTGCTTGTGTCTGTGCTGGGGACCACCACCATGCGTGTGGACATTCCTGCCC	2400
TGGGCGTGAGCGTCACCTTCAATGGCCAAGTCTTCCAGGCCCGGCTGCCCTACAGCCTCT	2460
TCCACAACAACACCGAGGGCCAGTGCGGCACCTGCACCAACAACCAGAGGGACGACTGTC	2520
TCCAGCGGGACGGAACCACTGCCGCCAGTTGCAAGGACATGGCCAAGACGTGGCTGGTCC	2580
CCGACAGCAGAAAGGATGGCTGCTGGCCCCGACTGGCACACCCCCCACTGCCAGCCCCG	2640
CAGCCCGGTGTCTAGCACACCCCCG 2669	

### REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

### Patent documents cited in the description

EP 96935358 A [0208]

4

### Non-patent literature cited in the description

- KNUDSEN, K.A. et al. Expl. Cell Res., 1990, vol. 188, 175-184 [0003]
- MERGE, R.M. et al. J. Cell Sci., 1992, vol. 103, 897-906 [0003] [0003]
- DONALIES, M. et al. Proc. Natl. Acad. Sci., 1991, vol. 88, 8024-8028 [0003]
- ROSEN, G.D. et al. Cell, 1992, vol. 69, 1107-1119 [0003]
- MORRISON, T.G. Virus Res., 1988, vol. 10, 113-136 [0005]
- BIOBEL, C.P. et al. Nature, 1992, vol. 356, 248-252 [0005]
- a guide to method and application. MICHAEL A.I. et al. PCR Protocols. Academic Press, 1990 [0035]
- SAMBROOK J. et al. Molecular Cloning a Laboratory Manual. Cold Spring Harbor Laboratory, 1989
  [0038]
- STANLEYT.; CROOKE; BERNALD LEBLEU. Antisense Research and Applications. CRC Publishing, 1993 [0044]

- BORREBAECK, C.A.K. J. Immunol., Meth., 1989, vol. 123, 157 [0064]
- TOSHIO KUDO. Tissue Culture, 1993, vol. 19, 61-65 [0064]
- CARTER et al. Pro. Nat. Acad. Sci, 1992, vol. 89, 4285 [0064]
- Molecular Cloning, a Laboratory Manual. Cold Spring Harbor Laboratory, 1989 [0069]
- SAMBROOK J. et al. Molecular Cloning, a Laboratory Manual. Cold Spring Harbor Laboratory, 1989 [0093]
- E. HARLOW; D.LANE et al. Antibodies, A Laboratory Manual. Cold Spring Harbor Laboratory [0093]
- CHOMCZYNSKI P.; SACCHI N. Anal. Biochem., 1987, vol. 162, 156-159 [0094]
- SMITH, D.B.; JOHNSON, K.S. Gene, 1988, vol. 67, 31-40 [0104]
- Molecular Cloning: A Laboratory Manual. Cold Spring Habor Laboratory, 1989 [0116]
- MIZUSHIMA S.; NAGATA S. Nucleic Acid Res., 1990, vol. 18, 5322 [0121]